Technical Report 346

Hierarchical Shape Description of Objects by Selection and Modification of Prototypes

John M. Hollerbach

MIT Artificial Intelligence Laboratory

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643.

HIERARCHICAL SHAPE DESCRIPTION OF OBJECTS BY SELECTION AND MODIFICATION OF PROTOTYPES

by

John M. Hollerbach

Massachusetts Institute of Technology

November, 1975

Revised version of a dissertation submitted to the Department of Electrical Engineering on January 24, 1975 in partial fulfillment for the degree of Master of Science.

Abstract

An approach towards shape description, based on prototype modification and generalized cylinders, has been developed and applied to the object domains pottery and polyhedra:

1. A program describes and identifies pottery from vase outlines entered as lists of points. The descriptions have been modeled after descriptions by archeologists, with the result that identifications made by the program are remarkably consistent with those of the archeologists. It has been possible to quantify their shape descriptors, which are everyday terms in our language applied to many sorts of objects besides pottery, so that the resulting descriptions seem very natural.

2. New parsing strategies for polyhedra overcome some limitations of previous work. A special feature is that the processes of parsing and identification are carried out simultaneously.

With this descriptive approach, the evidently unrelated domains of pottery and polyhedra are treated similarly. Objects are segmented into multiple generalized cylinders. The cylinders are then described by assigning a prototype, a standard shape from a small repertoire, which is modified to conform more exactly with the cylinder. The modifications are structured hierarchically and specify the degree of modification as coarsely or precisely as desired. Some modifications are specific to a given prototype, others are applicable to several of them.

The emphasis throughout this work has been to develop useful, qualitative descriptions which bring out the significant features and subordinate lesser ones. To this purpose curved lines representing the boundary of vases have been quantized into a few curvature levels. Line, region, and volume shapes are all described by assigning and modifying prototypes. In each instance the prototypes are specialized to the domain, and pose different problems in selection and modification.

Thesis Supervisor: Patrick H. Winston
Title: Associate Professor of Electrical Engineering and Computer Science

Acknowledgment

I would like to thank Patrick H. Winston for advising this thesis and for providing invaluable suggestions and assistance through the long course of this work. Eugene Freuder's knowledge and viewpoints about polyhedra helped mold my thinking in this area. Scott Fahlmann's early criticisms of my ideas aided their development. Many other members of the Vision Group and of the AI Laboratory, too numerous to mention, also provided useful suggestions along the way. Of these I should single out David Waltz, David Marr, and Richard Gabriel.

Table of Contents

THE SHAPE OF TRUTH

A Fable

A sage, who had filled his glass
     at the fountain of truth,
said, in a statement
     that later became canonical,
to his disciples,
     patterns of eager youth:
'I have seen truth itself;
     and it is conical'.

Piet Hein

CHAPTER 1 -- INTRODUCTION

Past work in machine vision has not resulted in a sound theory of shape description despite recognition that good description is prerequisite to the development of intelligent programs. This thesis attempts to ameliorate this deficiency.

HIERARCHICAL DESCRIPTION BY PROTOTYPE MODIFICATION

The idea behind the approach is that description should start with generalities and work toward specifics, that it is important to first have an overview before details are placed in perspective. An overview is established by assigning a prototype, a presumably simpler and more basic shape than the object being described. Since the prototype and the object ordinarily differ in exact details of shape, the prototype is modified in specific ways to conform more closely to the object shape. Prototype assignment and modification is similar to the schema and correction of Gombrich[1965], and exhibits some aspects of frame systems [Minsky 1974].

The modifiers go from a coarse to fine specification of the degree of modification. A width modifier might be quantized coarsely into the levels narrow and broad. If a finer differentiation of the width continuum is needed, each level may be split into sublevels, such as narrow into very narrow and slightly narrow. This process may be carried out to any level

of detail required. What is important is that the coarse description stands out, and that the finer detail is left unspecified unless required, which it often is not. Thus the various aspects of shape are described qualitatively rather than in terms of mathematically defined shapes.

As an example of mathematical versus qualitative description, consider curved lines. They have been approximated variously by straight lines, by circular arcs, and by polynomials. There are psychological objections to these mathematical approximations (enumerated in section 2.4.1), but their main failing is in rendering curvature too precise for recognition. A qualitative description, on the other hand, brings out general trends by quantizing curvature and by assigning labels to the quantum levels; a line is described, for example, as "strongly curved becoming gradually straight". With qualitative descriptions, higher level terms such as bow, hook, or stirrup shaped are readily assigned to lines.

Prototype assignment and modifier quantization induces a hierarchical description, whose merits are threefold:

1. Approximation is straightforward by disregarding lower levels of the hierarchy.

2. The higher levels can serve to index a description for modeling and identification.

3. The description can be made arbitrarily precise by adding depth to the hierarchy.

SOLID OBJECTS ARE MODELED AS GENERALIZED CYLINDERS

Past work in describing 3-D shape has failed to explicitly represent

the third dimension, and has proved inadequate for recognition as a result. Binford [1971] has recently put forth a solid object representation that does explicitly include the third dimension. His generalized cylinder scheme appears to be a fruitful one, and has already been applied towards shape description [Agin 1972, Hollerbach 1972, and Gabriel 1972]. The term ***generalized cylinder*** is derived from a generalization of an ordinary cylinder, which can be described as the movement of a circular cross section along a straight axis from one end of the cylinder to the other. The generalization consists of allowing the cross section and axis to assume arbitrary shape.

Generalized cylinders facilitate development of a hierarchical description. They induce a segmentation of an object into parts that are well described by prototypes with modifications. These parts, moreover, can be hierarchically arranged on the basis of size or significance. By placing certain restrictions on the formal definition of a generalized cylinder, the act of fitting a cylinder to a part has a smoothing effect that both provides a first-order approximation of shape and indicates where modifications are needed.

There are important differences in the implementation of the generalized cylinder concept in the above works. Binford's original formulation was extended and partially implemented by Agin [1972]. Recently Gabriel [1973] put forth his own version called suspensions. The latter two approaches are distinguished by their mathematically precise nature, as contrasted to the qualitative emphasis here. In addition,

although both Agin and Gabriel mention hierarchy, they do not present a method for achieving it.

DESCRIPTIONS OF POTTERY REFLECT ARCHEOLOGICAL USAGE

Besides hierarchy and generalized cylinders, the third important feature of this thesis is the correspondence of the approach with archeologists' descriptions of shape. Their descriptions appear to be hierarchical and can readily be placed into a generalized cylinder scheme.

The approach has been applied to two domains of objects: pottery and polyhedra. My study of description originally began with the polyhedral domain, from which the general approach evolved. The formal nature of this domain makes it particularly easy to apply the generalized cylinder concept. Later the study was particularized to understanding the shape of pottery in the terms normally used by archeologists. The advantages of the pottery domain are twofold: (1) there are numerous archeological books describing vases which can serve as a basis for study and comparison; and (2) it is a relatively simple yet sufficiently rich curved object domain.

Studying these books led to the conclusion that archeologists implicitly use the types of description advocated here. I read hundreds of descriptions of vases, noted which terms were used, and distilled the relationships among them. I found that the terms are on the whole applied precisely and consistently, not only across a single artheologist's descriptions, but across most of the archeologists whose books I read. This consistency allowed me to quantify many archeological terms -- terms

that are also common in everyday shape description. Descriptions of objects in this thesis therefore have a natural flavor.

To demonstrate the feasibility of the approach in the pottery domain, I wrote a program to describe and identify a vase starting from a list of points on its contour. The program first derives a qualitative description, and then uses this description to categorize the vase as one of 42 types. Some of the less familiar vase types recognized by the program are illustrated in figure 1.1.

The program describes only the main cylinder of the vase, which involves: (1) possible segmentation into foot, body, neck, and lip; (2) a description of each of these parts in terms of prototypes and modifications; and (3) the joining of the parts in a complete description. Handles or spouts are not described or segmented, although the terms in which this may be done are presented. The program structure and results are discussed in section 3.

To illustrate the types of description generated by this program and the sorts of terms used, the two vases in figure 1.2 are described below. For vase A:

vase type:	amphora, used for storing solids.
body:	tall ovoid, high-shouldered with straight lower profile becoming abruptly rounded.
neck:	high and broad cylinder, with straight and vertical profile, and offset from the body.
lip:	rolled.
foot:	low and narrow molded.
handles:	two vertical handles from shoulder to neck.

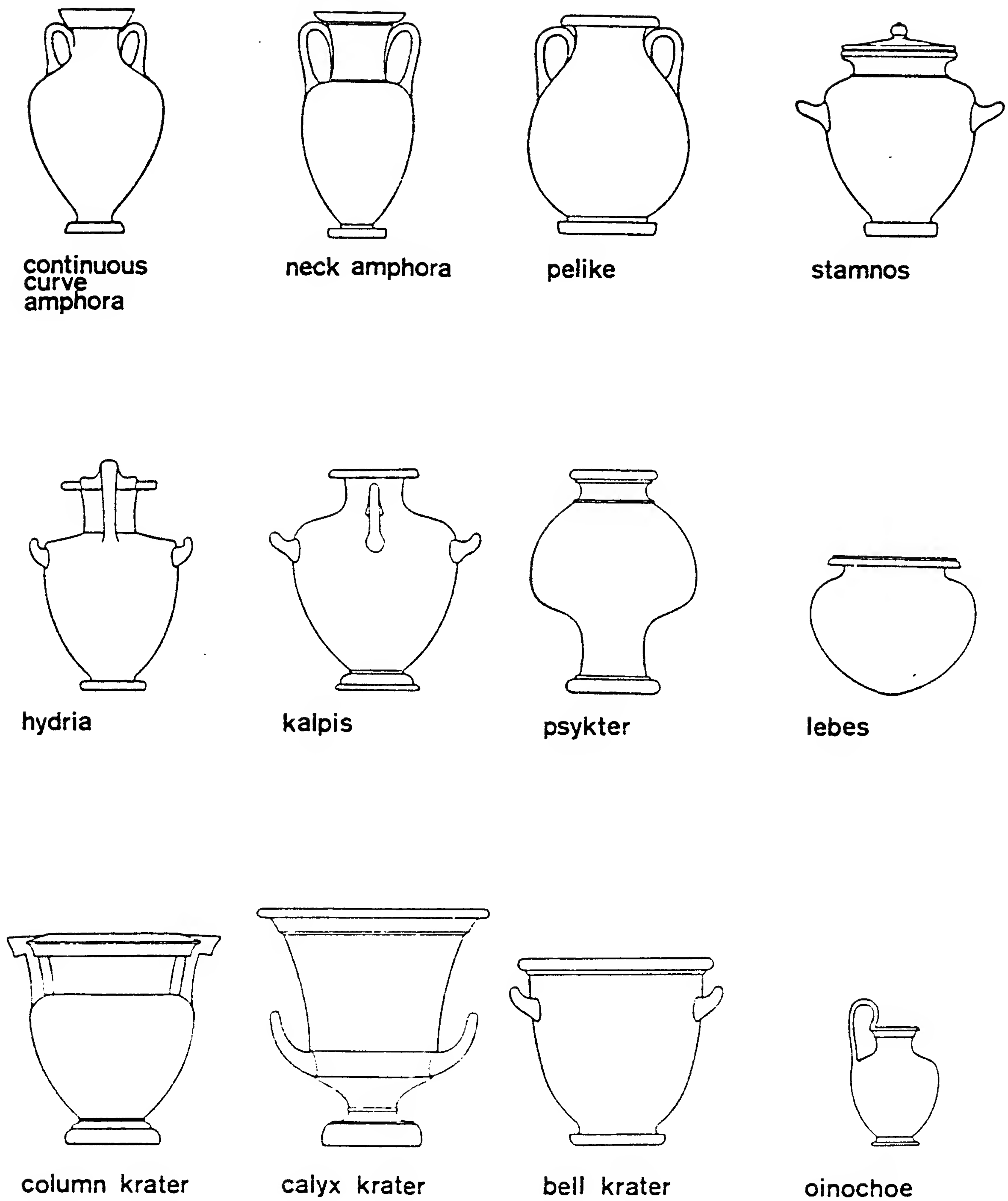


FIGURE 1.1. Examples of various Greek vases, from Cook (1960).

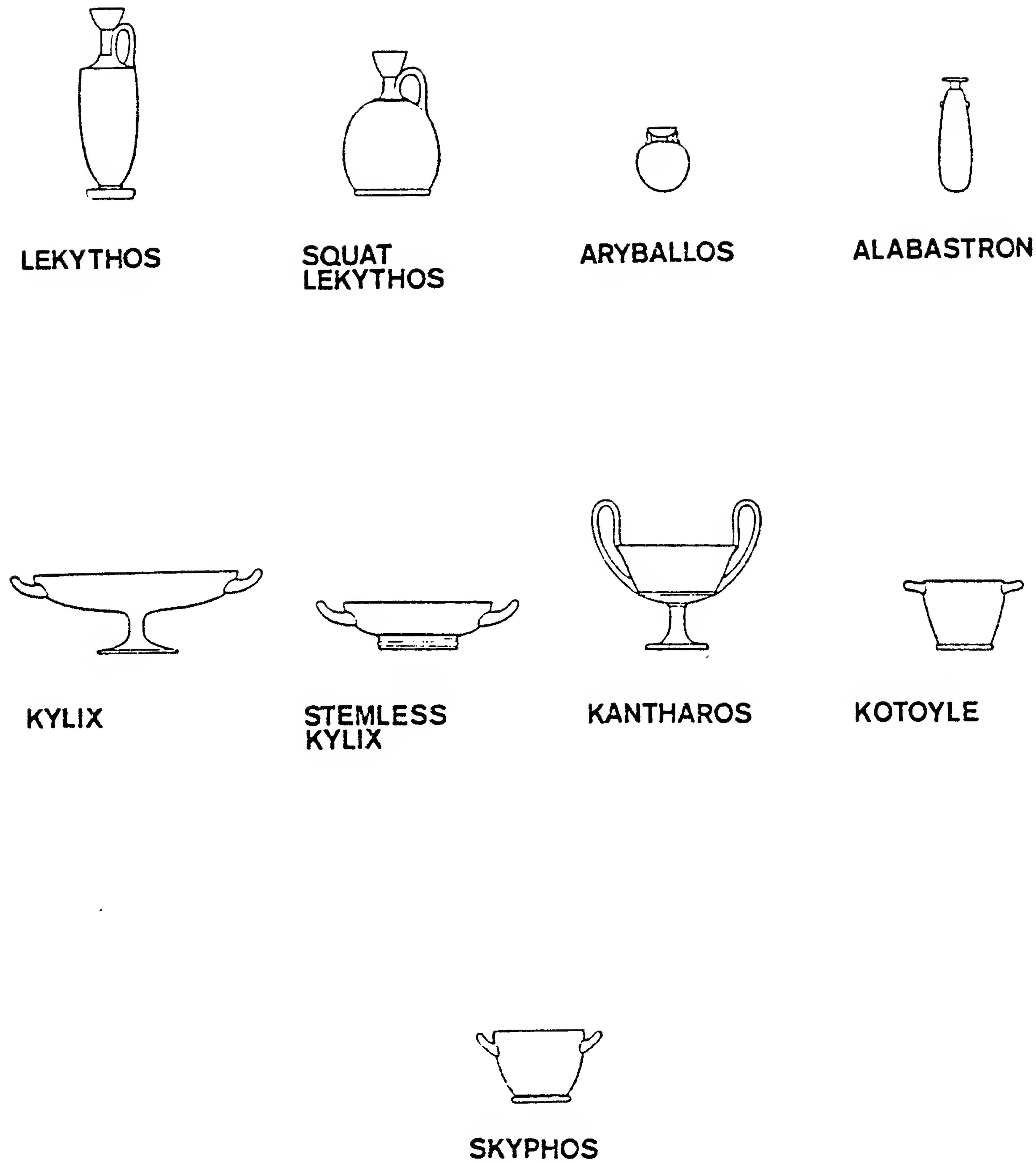


FIGURE 1.1. Examples of various Greek vases, from <u>Cook</u> (1960).

A. AMPHORA

B. KYLIX

FIGURE 1.2.

For vase B:

vase type:	kylix, used for pouring liquids.
body:	shallow bowl, open-mouthed with convex rounded profile.
lip:	very low molded.
foot:	high pedestal, widely splaying with broad stem and narrow base, and offset from the body.
handles:	two horizontal handles rising at a low angle from the body.

The general approach is presented in the context of both the polyhedral and pottery domains in section 2. The polyhedral domain is studied in greater detail in section 4. Section 5 presents conclusions and suggestions for further work.

CHAPTER 2 -- THE GENERAL APPROACH

2.1 Hierarchical Description

Visual recognition of objects requires the ability to set up some form of description of an object, to extract differences or similarities between object descriptions, and to rate these differences in terms of significance [Winston 1970]. A rating system implies not only that some differences are more important than others, but also that the descriptions are set up in such a way that comparison is meaningful. A hierarchical description can meet both of these requirements in a natural way.

As an example of such a description, consider object A in figure 2.1. If asked whether it is more like object B or C, we would probably choose B. Thus we have judged that the difference between A and B, namely the small indentation at the bottom, is less significant than the difference between A and C, namely the large indentation at the top. Size was evidently used as a comparative measure. If pressed further whether A is more like cube D or block E, again most of us would probably choose E. By so doing we have placed an interpretation on the top portion of the object. Rather than describing it as a cube with a top protrusion, we have judged it to be like a block with a top indentation.

The end result of these comparisons is a hierarchical description, where the level of a feature in the hierarchy is related to its importance.

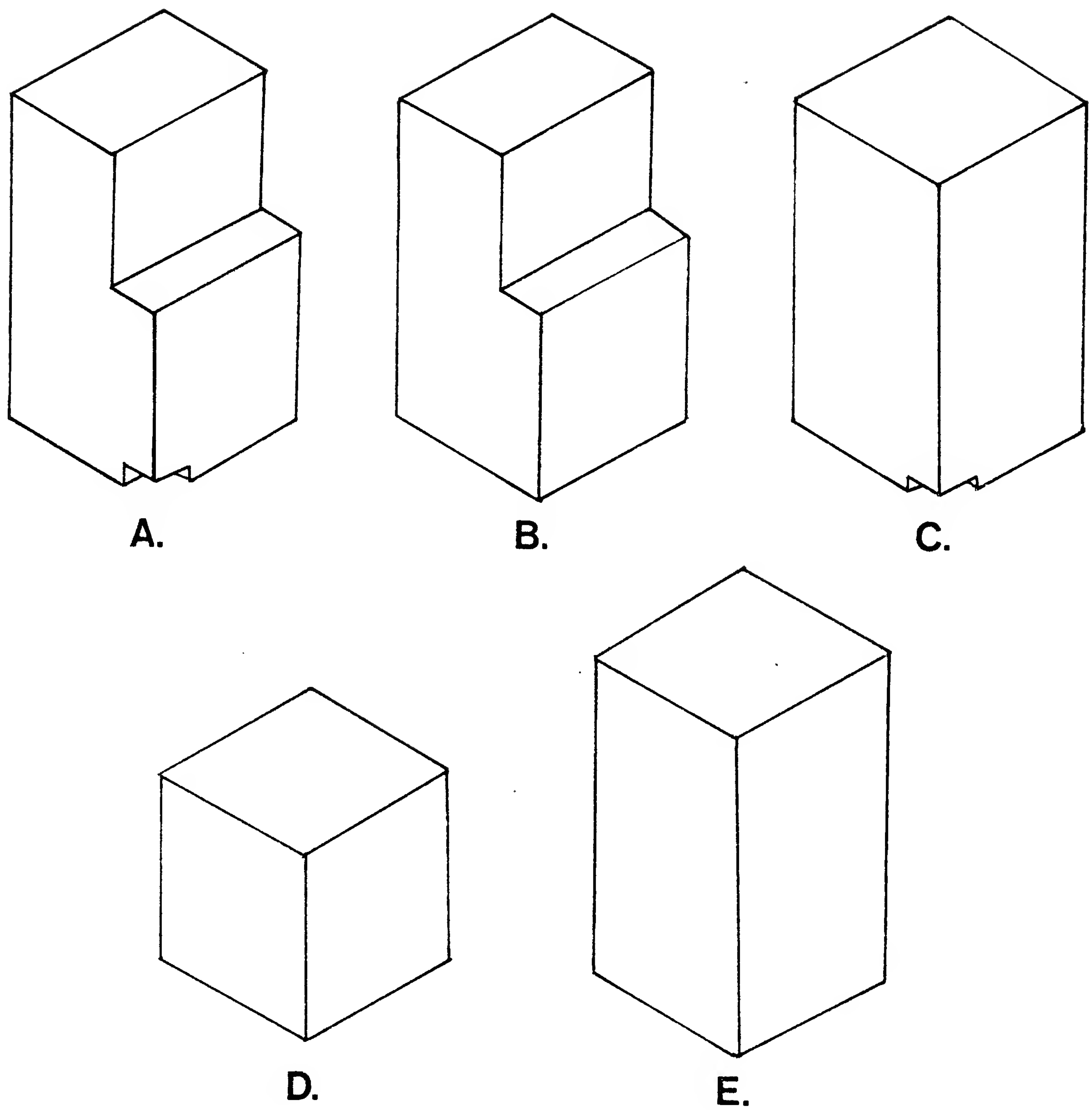


FIGURE 2.1. Object A is more similar to B than to C, and to E than to D.

Apparently the block likeness of A is the most significant feature, and so block is at the top level of the hierarchy. Next in importance is the top indentation, which occupies the second level. At the third level comes the bottom indentation, the least important feature.

Comparison between hierarchical descriptions is therefore conducted by matching levels from the top. The level at which a mismatch arises indicates the degree of similarity. Objects A, B, and C in figure 2.2 are all blocks at the top level, but at the second level A and B match while A and C do not. Therefore A is more similar to B.

ASSIGNING A FRAMEWORK PLACES DETAIL IN PERSPECTIVE

The process of assigning an approximate shape, such as *block* to A in figure 2.1, then modifying it hierarchically to conform more exactly with the object, establishes a framework for interpreting detail. It is only because A was placed in a block framework that the top and bottom indentations were interpreted as such rather than as something else.

The importance of placing detail into some larger framework is illustrated in figure 2.3A, where a window has been selectively placed on some portion of a vase. What does it represent, minor detail or significant feature? It can be either, as indicated by B or C. Any approach that seeks a description in a piecewise manner, namely by breaking the object into little pieces of contour and describing it as a collection of such pieces, would fail on just such an example. Without some overview, a piecewise approach is bound to become entangled in the weeds of

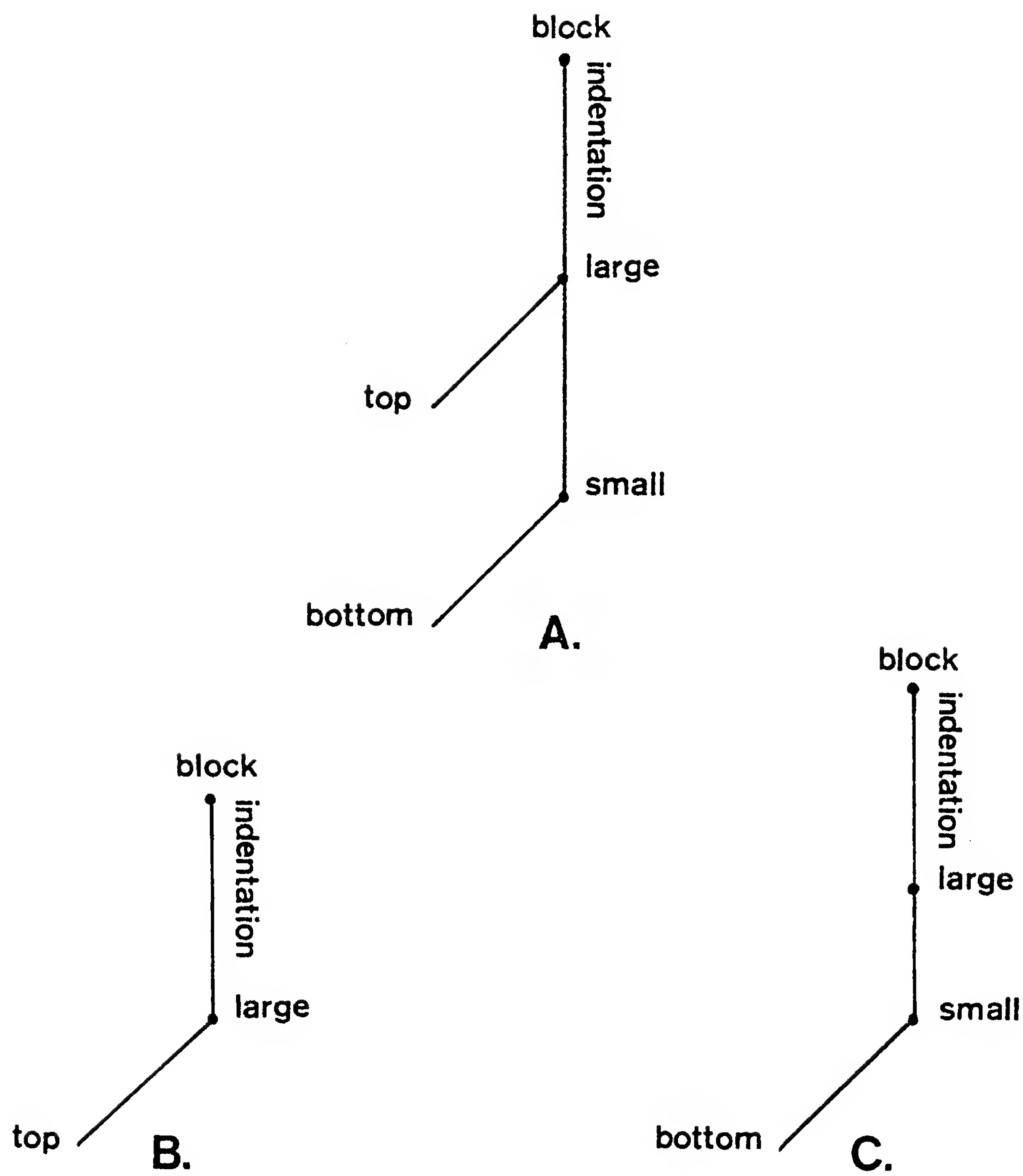


FIGURE 2.2. A comparison of the hierarchical descriptions of blocks A, B, and C would indicate A is closest to B.

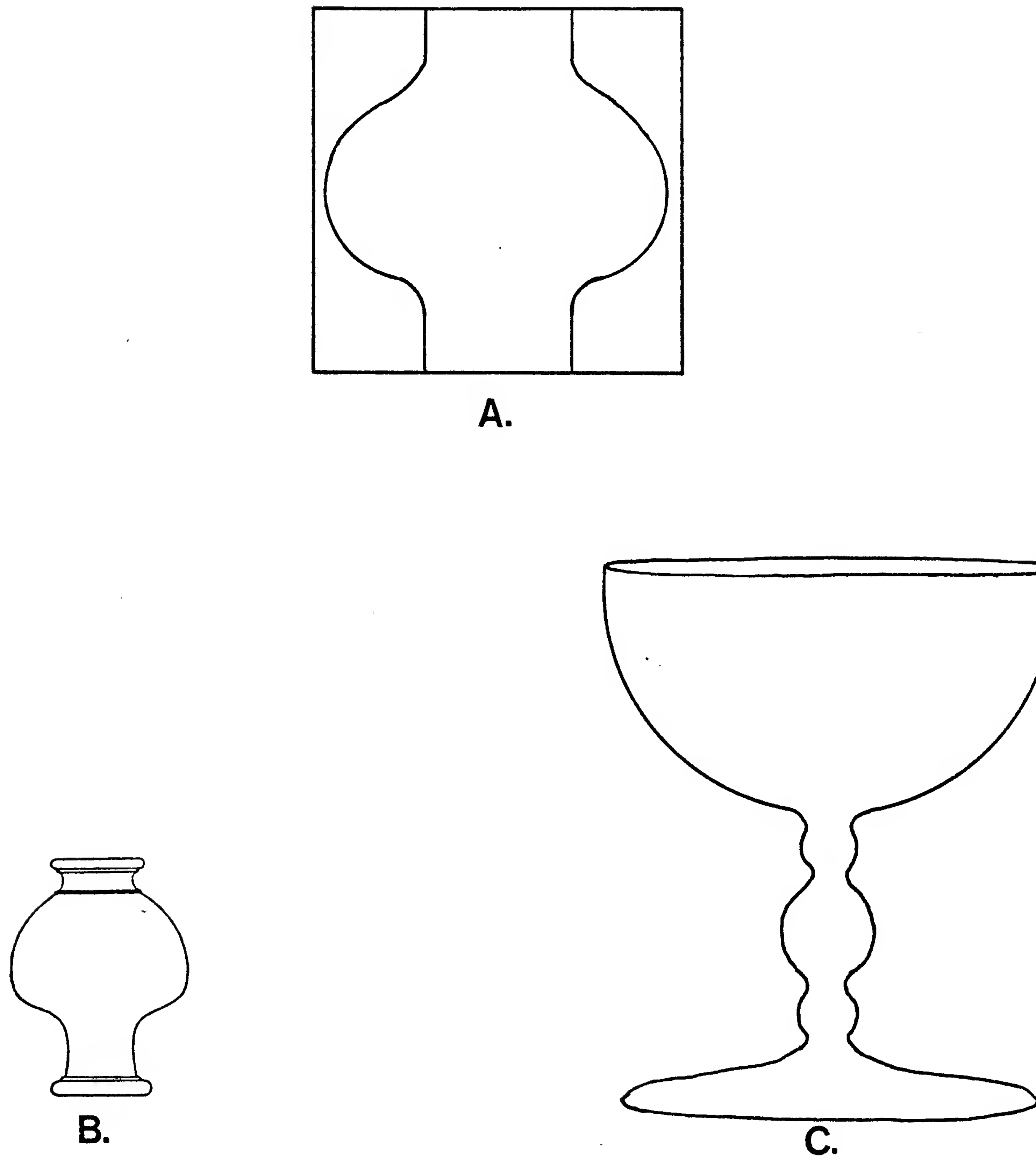


FIGURE 2.3. The selected window of a vase portion A can represent a significant feature as in B or a minor detail as in C.

irrelevant detail.

[Guzman 1967] is an example of a piecewise approach, whose limitations are illustrated by A in figure 2.4. Guzman describes A as a connected set of regions 1 through 7, where each region is represented by its boundary, namely as a concatenated set of straight lines. Comparison between objects A, B, and C is difficult with such a description. Which regions in A correspond to the three regions of B or C? If topological mappings are used, only regions 2, 3 and 4 of A could be matched against those of B or C.

A FIXED REPERTOIRE OF PROTOTYPES SERVES AS FRAMEWORKS

A number of approximate or rough shapes are needed to handle a wide variety of shapes. If there are too many of them, the basic similarities between objects may not be brought out. If too few, they may not correspond closely enough to possible object shapes. Since the number of approximate shapes will be much smaller than the number of possible shapes, some mismatch will arise which is diminished by modification.

The repertoire of approximate shapes can be considered a set of prototypes for object shapes. Simple shapes presumably make better prototypes than more complicated ones; for example, *block* is a better prototype than figure 2.1B or C. On the other hand, a complicated shape may be so common in a visual domain that it deserves its own prototype, such as bell-shaped in Western culture.

The exact nature of prototypes is not crucially important, as long as

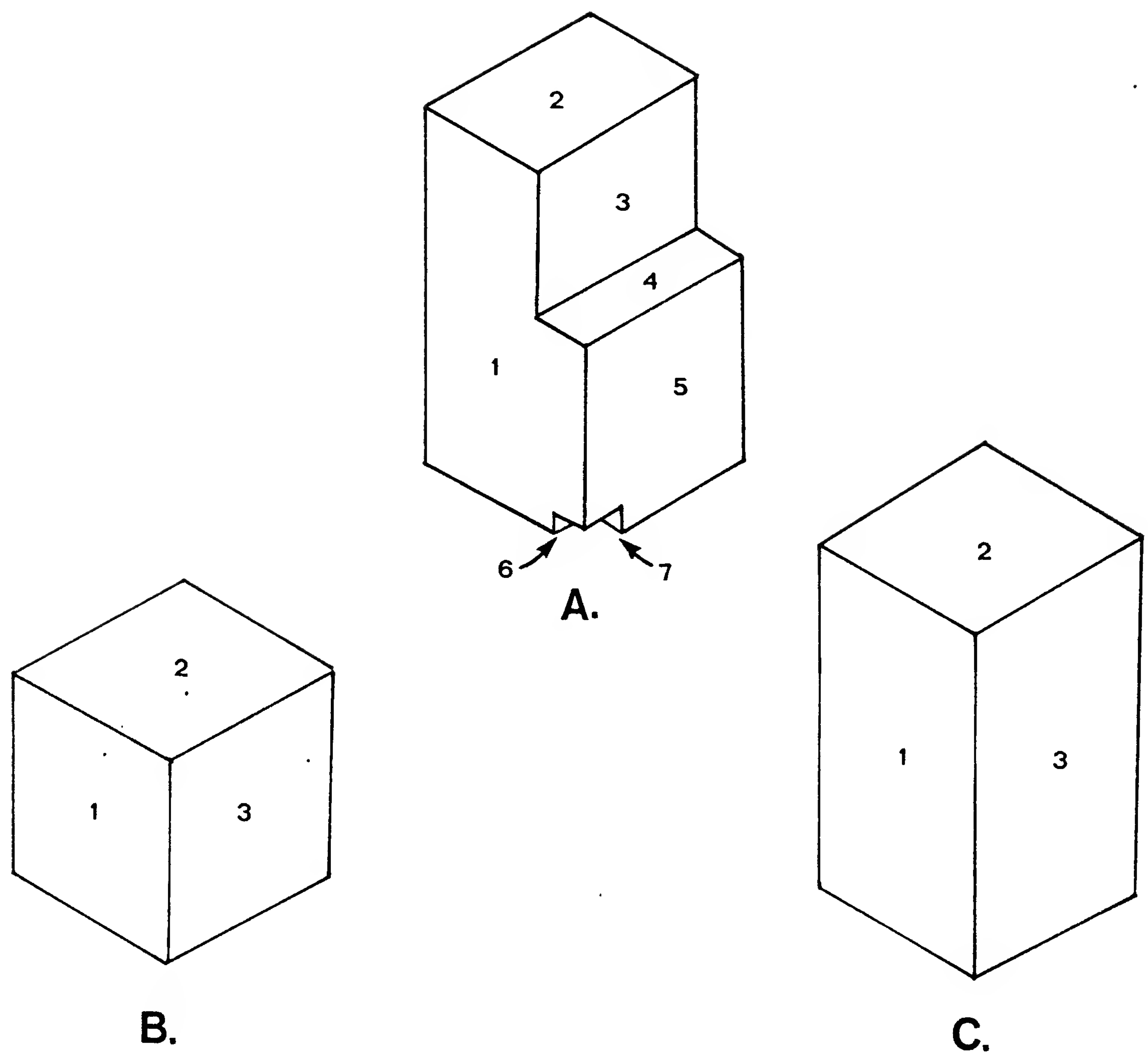


FIGURE 2.4. Topological modeling makes matching difficult. Which regions of A should be matched against the regions of B and C?

they adequately characterize common shapes. *Ellipsoid* and *ovoid* are fairly similar, and one can easily be described in terms of the other. Which is chosen is therefore somewhat arbitrary, presuming they are too similar to be both prototypes. Of course in certain domains one choice may be more appropriate than another. *Ovoid* is often used in describing pottery, evidently because of the frequent similarity of vase parts and eggs.

MODIFIERS ASSOCIATED WITH A PROTOTYPE IMPOSE AN INTERPRETATION TO DETAIL

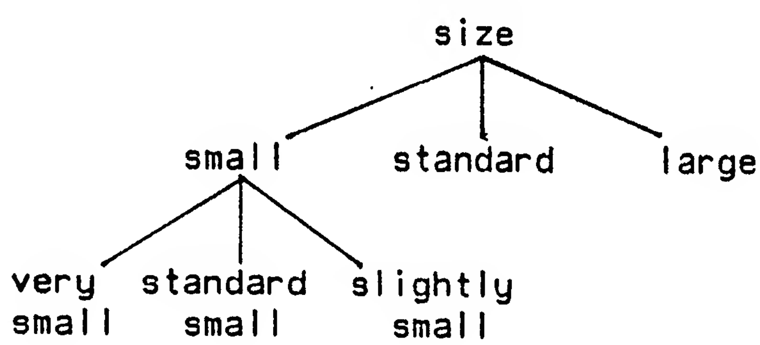
A given prototype has associated with it a number of ways in which it can be modified. Thus *block* may be modified by indentations or protrusions, while *ovoid* may be modified among other ways by altering the height-width ratio or the height of the shoulder. Some modifier types are general to a number of prototypes, others are specific to just one. Thus *cone* and *ovoid* both have a height-width modifier, whereas *cone* does not have a shoulder modifier.

The set of modifier types associated with a prototype force a particular interpretation on the features of an object, as if there were a preexisting framework with slots to be filled. To speak of an object as an *ovoid* is to commit oneself to talking about its shoulder. This concept of prestructured frameworks is of current importance in Artificial Intelligence [Minsky 1974].

MODIFIERS QUALITATIVELY DIVIDE THE CONTINUUM

When one of the indentations of object A was said to be more

significant than the other, qualitative size difference was the essential dimension of comparison. Size of course varies continuously. For a symbolic description to work, the size continuum must be split into a number of levels: for example into small, standard, and large, where some standard interval is chosen with respect to which small and large are measured. The choice of standard interval depends on the nature of the object; a standard elephant is different in size from a standard mouse. If greater refinement is required, the levels may be split into sublevels, such as small into very small, standard small, and slightly small.



The continuum can eventually be approached in this manner, but it is unlikely that more than one or two levels will be necessary for the ordinary processes of description.

Thus the general format for a modification is the following:

```
(prototype modifier-type modifier
                         submodifier
                         subsubmodifier
                         etc.)
```

The modifier-type indicates how the prototype is being modified; for example, a *block* may be modified by an indentation, an *ovoid* by changing its height-width ratio.

COMPARING HIERARCHICAL DESCRIPTIONS FOR SIMILARITY OFTEN REQUIRES OUTSIDE MEDIATION

Object comparison increases in difficulty with the number of modifications to the prototypes. For example, one cannot decide on the basis of the hierarchical descriptions alone whether A in figure 2.5 is closer to B or to C because the amount of mismatch is the same. The mismatch between A and B consists of an extra protrusion for A and an extra indentation for B, while the mismatch between A and C consists of an extra indentation for A and an extra protrusion for C. Because these various modifications are of the same approximate size, they occupy equal positions in the hierarchies. Thus the mismatch between A and B is equivalent to that between A and C.

Disparity in the type of modifications is harder to reconcile than disparity within a particular type. Winston [1970] has addressed himself to this general problem. One of his suggestions, offered as possible but probably unsatisfactory, is a numerical rating scheme. Whereas some modifier types can be measured in the same way and therefore have equal significance, for example size used to measure indentations and protrusions, others cannot be compared directly, such as the height-width and shoulder height modifiers for ***ovoid***. The height-width modifier type may be considered to be more significant than the shoulder height modifier, and therefore would receive a higher numerical rating.

As Freiling [1973] has pointed out, symmetric matching where each object has equal weight might be useful for a few applications such as

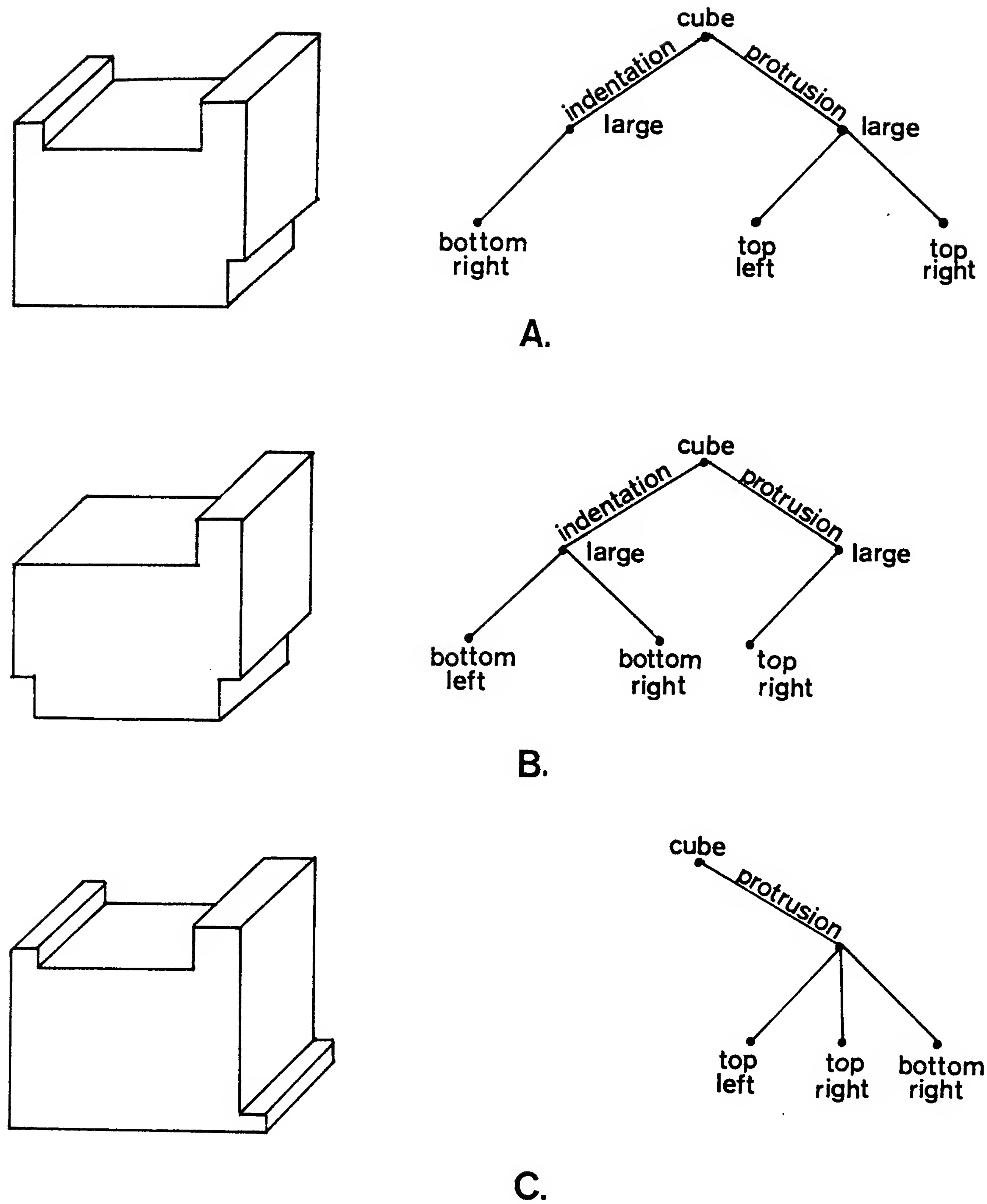


FIGURE 2.5. Similar objects with multiple modifications are difficult to rate in a pairwise comparison.

analogy problems, but for recognition purposes one really wants an asymmetric scheme. In judging whether a vase is an amphora, the requirements of similarity are explicitly mentioned in the model for amphora. The model may require the body to be a tall ovoid, while it allows the lip to range in shape as long as it is not too large. The model itself sets forth the conditions for matching, and this breaks the bind of reconciling different types of mismatch in a symmetrical matching scheme. One is almost forced into a numerical scheme for symmetric match because differences have to be rated over all possible object comparisons, independent of identity of any object.

SEGMENTATION OBVIATES THE NEED FOR MORE COMPLEX PROTOTYPES

Some objects may be too complex to describe with simple prototypes. Rather than create complicated prototypes for such objects, descriptive economy suggests segmenting them into simpler subparts more amenable to simple prototypes. A vase, for example, is ordinarily segmented into handles, foot, body, neck and lip, each of which can normally be assigned a simple prototype.

OBJECT SUBPARTS ARE RANKED HIERARCHICALLY BY SIZE

Size is the most generally useful criterion for ranking such subparts in a hierarchy, but functional importance may also play a role. Thus the keyhole of the padlock in figure 2.6 is functionally integral to identity. Although the keyhole is the same size in terms of visible area as the chunk

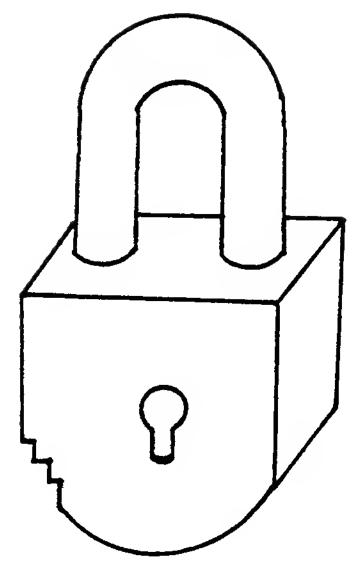


FIGURE 2.6. Significance can depend on more than just size, as a comparison of keyhole with a chunk missing form the side of the padlock reveals.

missing from the side, it is a more important part of the description. Note that functional importance cannot be determined without having an idea of probable identity. However, size can lead to the first coarse description from which identity can be postulated. This subject is pursued further in section 2.3.2.

The body of a vase is the largest subpart, and its shape is important in determining identity. The lip on the other hand is often the smallest subpart, and its shape within fairly broad limits is relatively unimportant. Since body shape is so important, presumably it would need to be more exactly known than lip shape. This relation holds in general, and is restated below:

1. The more significant a part is in some ranking, the more detailed must be its description.

2. Conversely, the lower a part is ranked, the less detailed is its description.

3. Size is an important ranking criterion.

2.2 Generalized Cylinders

The basic ingredients for a generalized cylinder are an axis and a cross section that moves along this axis (figure 2.7). The cross section is required to lie in a plane that is perpendicular to the axis throughout the movement. There are a number of ways in which a generalized cylinder can be defined more precisely, depending on what restrictions are placed on the axis and on the variability and manner of movement of the cross section. In its most general form, the axis is an arbitrary space curve while the cross section may freely change shape as it translates along the axis. A more constrained definition has been found useful in this thesis and consists of the following:

1. an axis lying in a plane
2. an arbitrary cross section with fixed shape
3. a continuous scale change function for the cross section as it moves along the axis

More generality in the definition is unnecessary for the types of objects under consideration here. To be sure, restriction to a plane means that space curves cannot adequately be described with just the shape descriptors proposed here. Nevertheless the number of objects that are describable with a 2-dimensional axis is large, and for those that cannot be so described the 2-dimensional case may be a good approximation.

TWO IMPLICATIONS OF THIS CONSTRAINED DEFINITION

Strictly speaking, a fixed shape cross section makes it impossible to

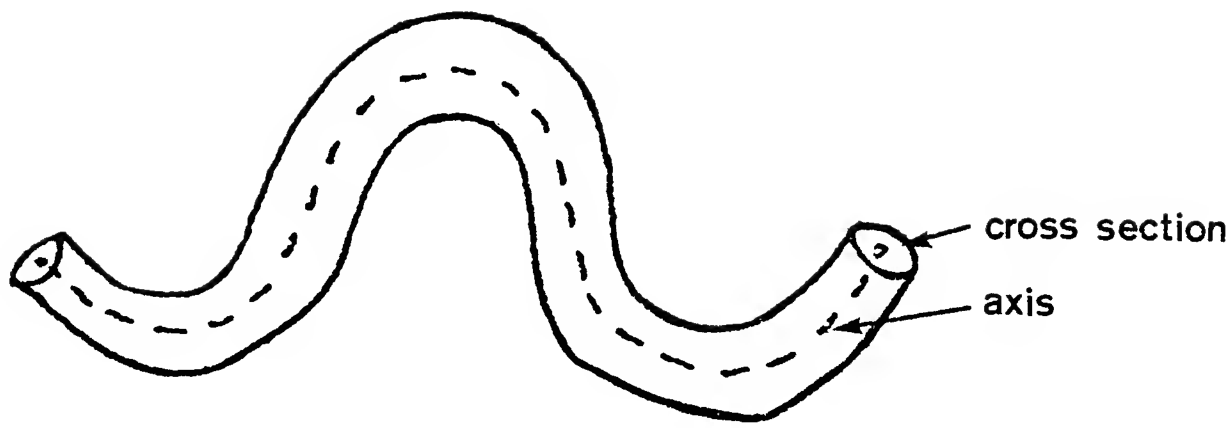


FIGURE 2.7. A generalized cylinder.

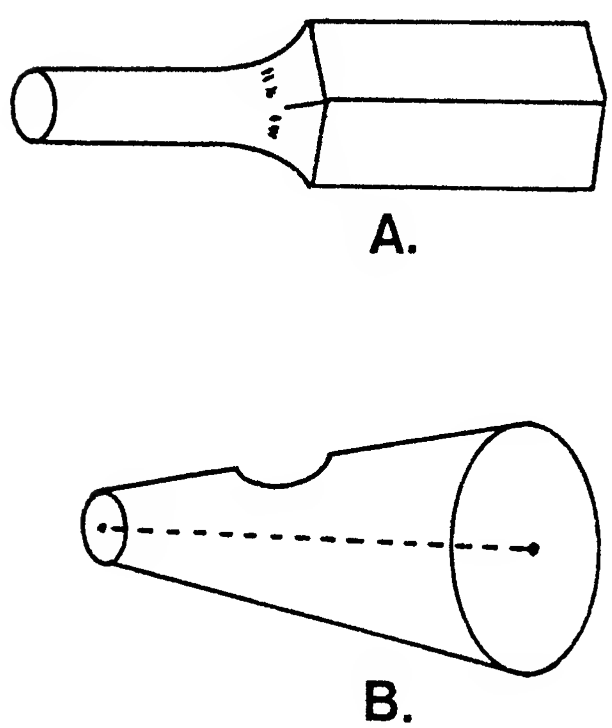


FIGURE 2.8.

model such objects as in figure 2.8; for example, A has a cross section varying drastically in shape. I have chosen not to relax restrictions to handle such cases, but prefer instead to (1) segment such objects at the point where the cross section changes grossly in shape, and (2) indicate qualitatively how the transition between the resulting parts occurs: articulated, smoothly, etc. I have not carried out a clearer specification of when to segment into distinct cylinders, and am mostly concerned with describing single cylinders in this thesis.

Scale change is also unable to account for the slight modification to the cylinder in figure 2.8B. A cross section changing only in scale is a good first order approximation because it smooths what might otherwise be an irregular cylinder. Any variations from continuous scale change can then be described as modifications to the smoothed cylinder; for example, B would be described as a cone with indentation.

2.2.1 The Appeal of Generalized Cylinders

Agin [1972] has discussed the intuitive appeal of generalized cylinders (p. 5):

> Many natural and manmade objects possess elongation.
> Most higher orders of life are distinguished by their
> extremities--legs, arms, heads, stems, and branches...
> And where elongation is present, the direction of elongation
> usually bears some useful or functional relationship
> to the object as a whole.

Thus Agin feels the axis of a generalized cylinder captures the general shape and orientation of elongated objects. The stick man (figure 2.9A),

for example, seems to capture the essential aspects of the human shape. Having a cross section move along the axis is like putting meat on the bones (figure 2.9B), and gives directly a three-dimensional or volume representation of objects.

CHILDREN DRAW OBJECTS AS GENERALIZED CYLINDERS

Generalized cylinders have more than just intuitive appeal, according to the experiments of Gluchoff [1973] with children. Children apparently conceptualize objects in a manner that is close to Gabriel's formulation of generalized cylinders. He represents cylinders as Susp(D1,D2), where D1 is one region, D2 an opposite one, and Susp a filling of the middle "in the simplest manner possible". For example (figure 2.10), a block is Susp(rectangle,rectangle) while a cone is Susp(point,circle).

A child might draw a wedge by connecting two triangles with lines (figure 2.11A). Similarly, a cylinder is drawn as two circles connected by two lines (figure 2.11B). Gluchoff's interpretation of these and similar experimental results is that children represent such objects by beginning and end faces and a filler in between (figure 2.12) -- analogous to Gabriel's formulation.

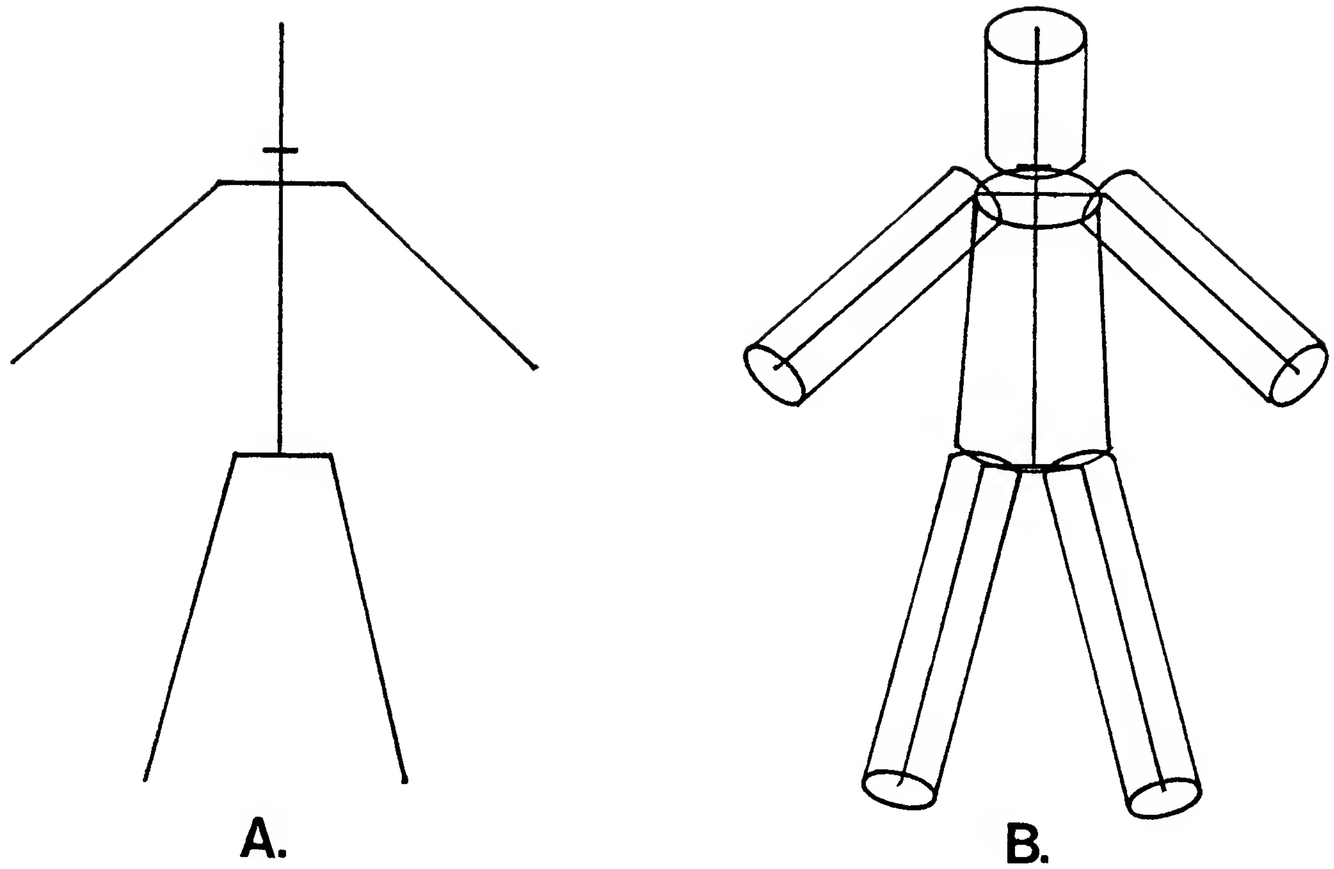


FIGURE 2.9. The stick man A, taken from Agin (1972), has meat on his bones in B.

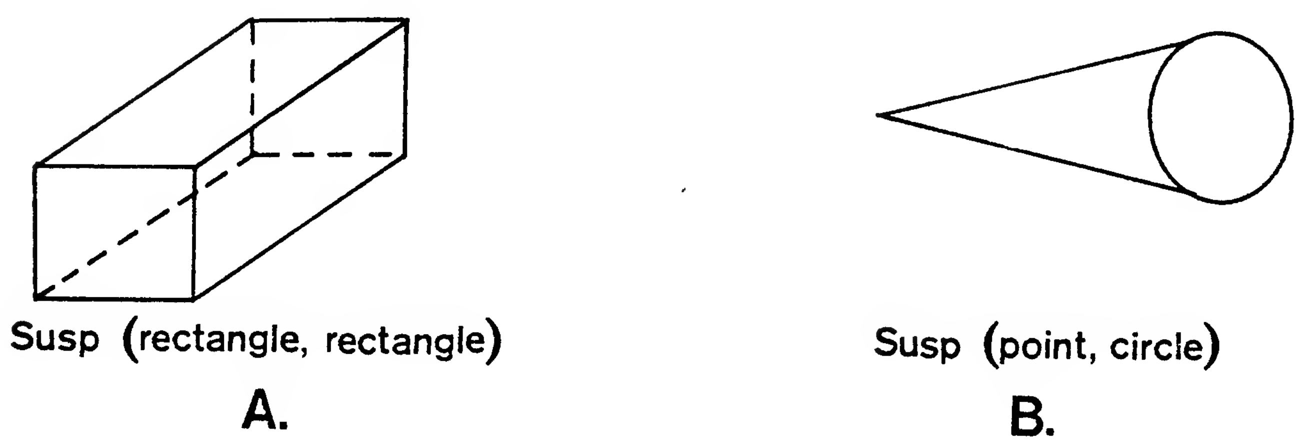


FIGURE 2.10. Examples of suspensions (Gabriel, 1973).

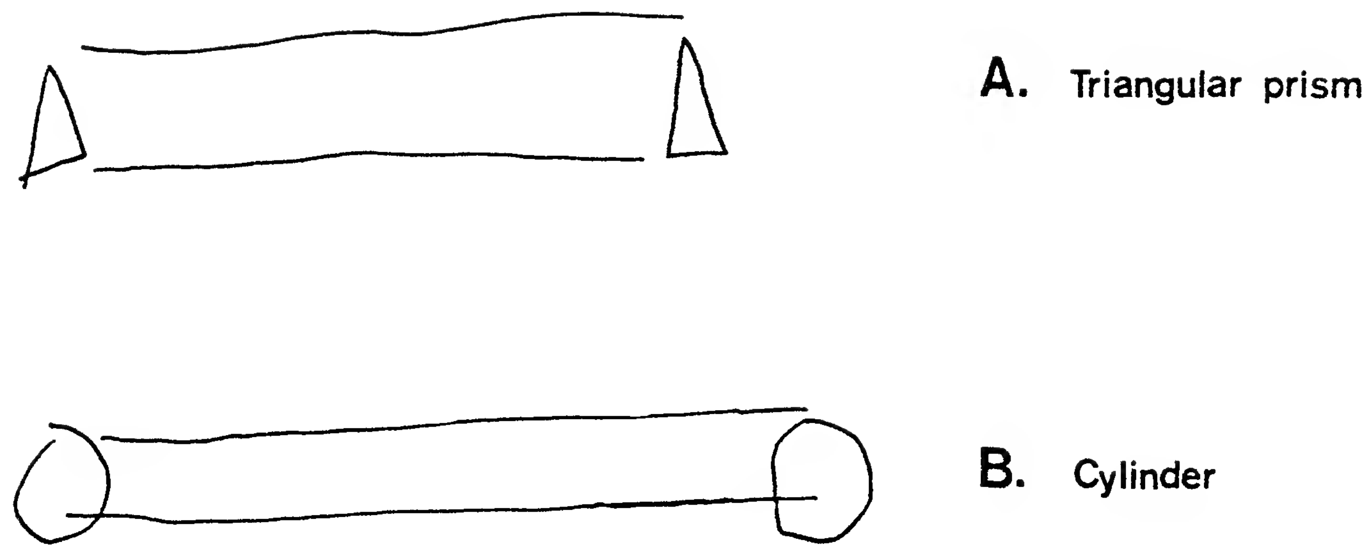


FIGURE 2.11. Some children's drawings from Gluckoff (1973) illustrating similarity with Gabriel's suspensions.

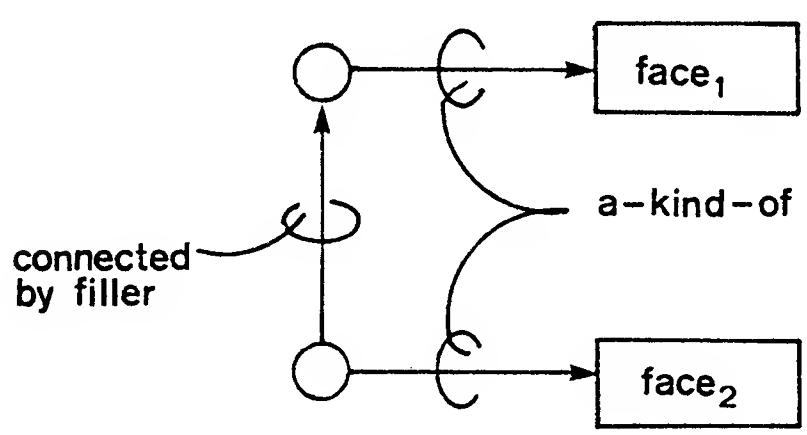


FIGURE 2.12. Structural diagram of a child's representation of an object from Gluckoff (1973).

2.3 Segmentation

This section presents a way of segmenting cylinders by examining their contours.

2.3.1 Contour as an Indicator of Shape

Of the various visual properties of objects that could be used for recognition, shape is probably the most important [Attneave 1967]. Other properties such as highlight and texture might also be fruitfully employed, as shown, for example, in [Krakauer 1971]. Nevertheless, this thesis confines the problem of recognition to that of developing reasonable shape descriptions.

Reasonable shape descriptions are most easily developed from object contour. Especially the rotational symmetry of vases makes contour a more attractive choice than such other shape indicators as texture, shading (see [Horn 1970]), or binocular disparity. Whatever way shape is derived, the types of shape description needed for recognition will likely be similar to those advocated in this thesis.

CONTOUR AND INTERNAL FEATURES

To determine the exact surface shape within the contour lines, one would have to examine internal features. Using internal features to predict shape is, however, difficult and potentially misleading. It is difficult, for example, to distinguish vase decorations from shape features. It is potentially misleading, as demonstrated in Agin [1972], to

segment cylinders by linking internal points; Agin's segmentation points are often badly placed.

Waltz [1972] and Shirai [1972] have demonstrated that, at least in the polyhedral domain, working from the outline inwards places the most constraints on scene and image analysis. Similarly, contour is the best guide for cylinder segmentation. For the purposes of this segmentation, one might as well assume that the cylinder surfaces are rounded (of circular cross section) -- an assumption that is obviously justified for pottery, and that is evidently used by humans as a default condition whenever curved outlines are perceived [Arnheim 1954]. After segmentation, it would then be appropriate to modify the assumed roundedness of individual cylinders by examining internal features.

2.3.2 Use of Contour for Segmentation

The problem of segmentation by contour is to pick out the major parts, given that the profile can vary wildly. Some of these variations may represent minor detail, others might signal a point of segmentation. A way to discriminate them is to start with a rough segmentation by applying general rules (discussed below). If the segmentation leads to a satisfactory description of the parts and of the whole, it is assumed correct. If not, the reason for failure is examined to decide on an alternate. After the new suggestion is applied, the process of creating a description is repeated.

This presumes the ability to judge what is and is not a satisfactory

description. To some extent it can be done on the basis of descriptive economy, i.e., perhaps the description is too complex and a simpler one could be obtained with an alternate segmentation. Domain specific rules take precedence over the general ones whenever they conflict.

LARGE SCALE CHANGE INDICATES A POSSIBLE SEGMENTATION POINT

The junction of two differently sized parts yields a change in scale. When the neck and body of a vase come together, for example, the scale changes dramatically from the relatively narrow neck to the broad confines of the body. Without such a change, the two-part configuration would probably look indivisible.

What is a large scale change, and how is it measured? The axis is divided into intervals, and for each interval the difference between maximum and minimum scale value is computed. Those intervals with scale change substantially above some threshold, such as the average, are selected as possible segmentation points.

The right choice of interval is important. If too small, minor variations in contour may yield large scale change locally and confuse the segmentation process. Too large an interval will diffuse the outline and cause possible segmentation points to be missed. Some intermediate choice is needed, a choice that reveals segmentation points while having a useful defocusing effect on the contour.

DOMAIN SPECIFIC KNOWLEDGE SELECTS THE RELEVANT LARGE SCALE CHANGES

Often several large scale changes are found for a given cylinder. To determine which represent appropriate points for segmentation, domain specific knowledge must be brought to bear. Thus knowledge that a vase ordinarily consists of a body, foot, neck and lip, and that these parts are related in certain ways, allows the scale changes to be interpreted more meaningfully.

Domain specific knowledge for pottery includes the following. The body is the largest part, and normally has a fairly smooth contour. The foot and neck tend to range in size from very small to a little more than half the body size. Junction with the body is ordinarily clearly delineated. A foot may be ornamented, which often leads to large changes in scale. The neck contour is almost always a simple curve. A lip may crown the neck, or be directly attached to the body. Lips do not normally reach a very large size, and may have an indistinct junction with the body or neck.

A segmentation strategy can be devised from this vase framework. Working from the bottom of the vase, the highest large scale change that yields a subpart of less than 30% area is called the foot. Working similarly from the top, the lowest large scale change that yields a subpart of less than 30% area is the neck assembly.

Thus all large scale changes except two can be ignored. The ones below the foot segmentation point are assumed to represent foot features, those above the neck segmentation point are neck or lip features. Any

large scale changes in the region 30% above the bottom and 30% below the top are normally assumed to represent body features.

For the handleless krater in figure 2.13, the four large scale changes a through d are found. Points a and b both yield a subpart less than 30% in area, and so b, the highest of these, is chosen as the foot segmentation point. Similarly, c is chosen as neck segmentation point even though d also yields a subpart of less than 30% area.

THE RATE OF CHANGE OF SCALE DETECTS SMALL LIPS AND FEET

For small lips and feet, the junction with the body is often too indistinct to be signalled by large scale change, as for the carinated bowl in figure 2.14. A more sensitive parameter, the rate of change of scale, is required for this circumstance. This parameter corresponds to change of curvature, and tends to amplify contour variations.

As pointed out in [Birkhoff 1933], people like to see gradual changes in curvature. Since gradual curvature is pleasing, sharp curvature is displeasing and attracts attention. Attneave [1954] conducted experiments in which subjects were asked to select the most representative points of various curved lines, and found that points of greatest change in curvature were chosen. Since such points are the most noticeable, they are also good segmentation points.

Sharp curvature may draw attention to body features as well as to foot and lip junctions. Thus the carination point of the bowl is as significant as the lip and foot junctions. Once again, domain dependent knowledge

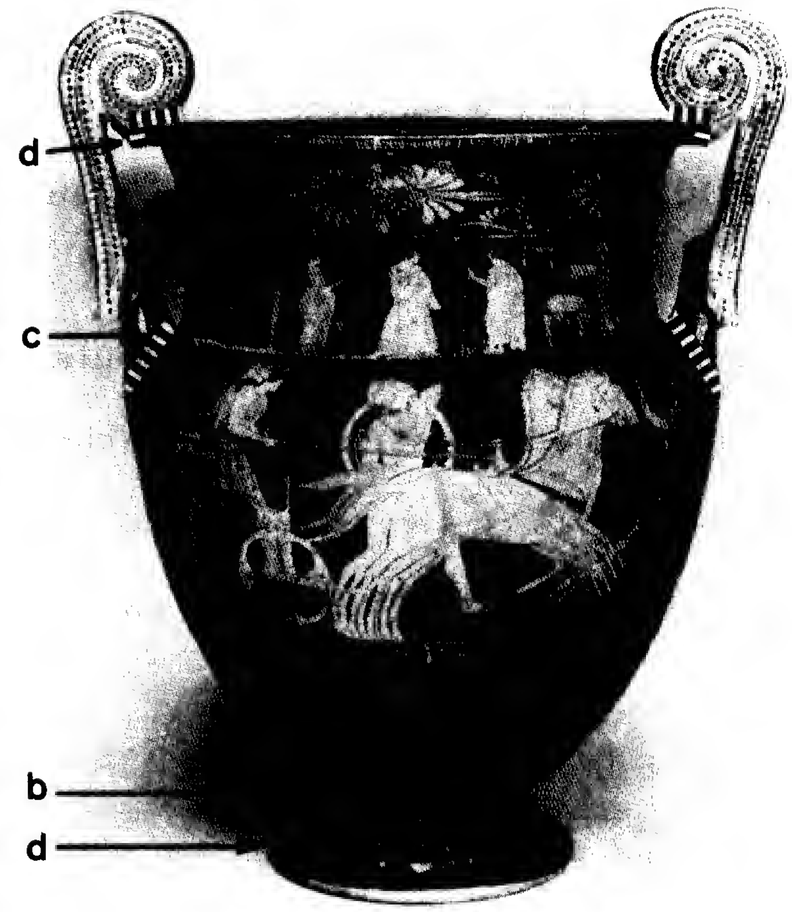


FIGURE 2.13. Four large scale changes at points (a,b,c,d) are found for this handleless krater.

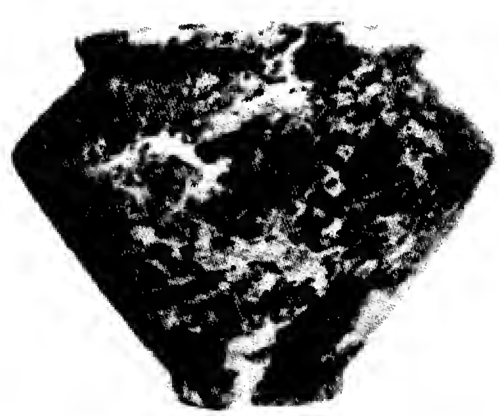


FIGURE 2.14. The small lip and foot of this carinated bowl are found by means of second width changes.

allows the lip and foot junction points to be accepted and the body point to be rejected. Small lips and feet missed with large scale change are estimated to comprise no more than 20% of the area. This rules out the body carination point, which would yield a subpart of larger than 20% area.

Gardin [1967] has pointed out that convention determines when the vase in figure 2.15 ceases being a concave body, as vases a or b, and becomes a convex body with a concave neck or lip, as vases d or e. The 20% value for small lips and feet yields this distinction, and pinpoints the border case c as a concave body.

AFTER SEGMENTATION, THE PROTOTYPE ASSIGNER CHECKS THE SUBPARTS

If the available prototypes require excessive modification to fit the segmented parts, a complaint is made and a different segmentation suggested. Because vase bodies are normally convex, the first segmentation usually results in a successful prototype assignment. Hence a complicated suggestion-verification process is not needed.

An hour glass shape, for one, causes the program to reject a segmentation and propose an alternate (figure 2.16). This particular shape may reach the prototype assigner if a neck and foot were successfully found above and below it. None of the available prototypes fits well enough, so the program suggests b as segmentation point. Note that altering the class of prototypes alters what does and does not fit. If there were an hour-glass prototype, the segmentation in question would not have failed at this point.

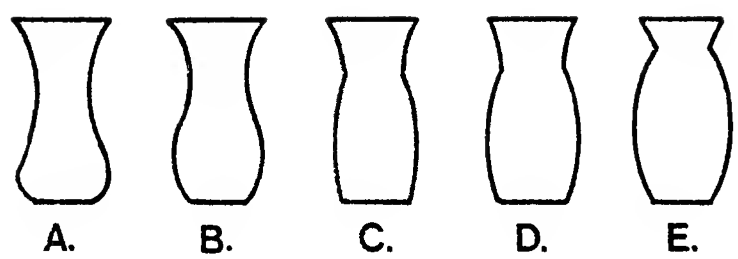


FIGURE 2.15. Variations in the delimitation of neck and body, from Gardin (1967).

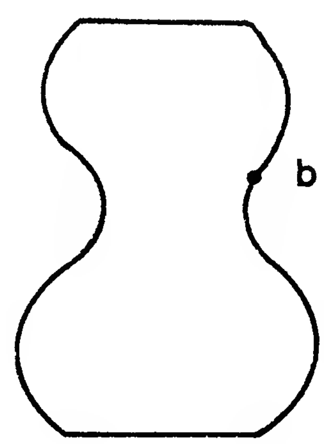


FIGURE 2.16. An hour glass shape is segmented further at point b.

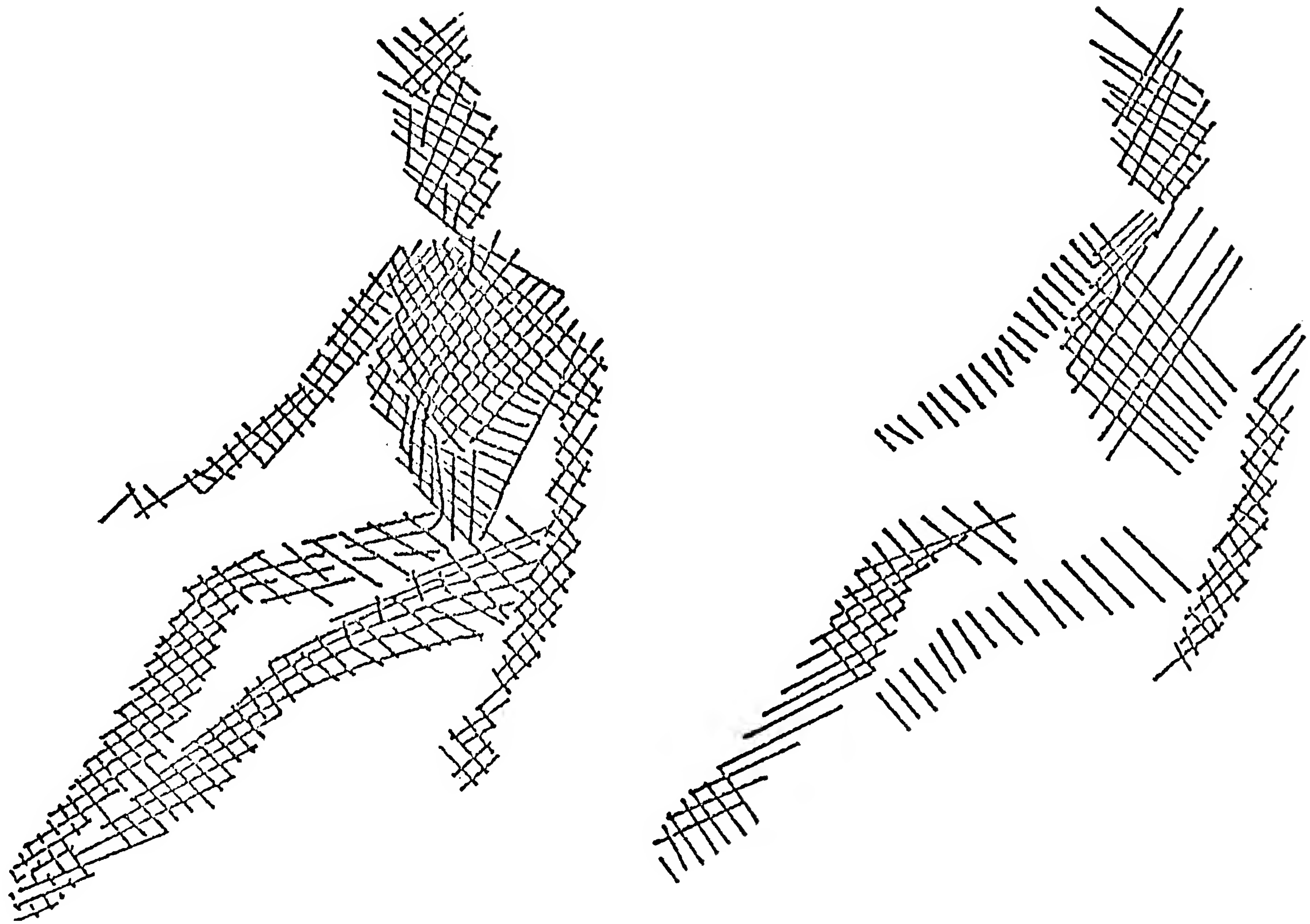
OTHER WORK ON CYLINDER SEGMENTATION

In this section the emphasis has been on segmenting single cylinders whose axes and orientations are known. The more general problem of segmenting into multiple cylinders and of estimating axis and orientation has been addressed in [Agin 1972] and in [Nevatia 1974].

Agin applies his approach to a set of intensity points obtained by laser scanning of an object resting against a dark background. The scanning is done with a plane instead of a line, so that one position of the plane results in a line reading.

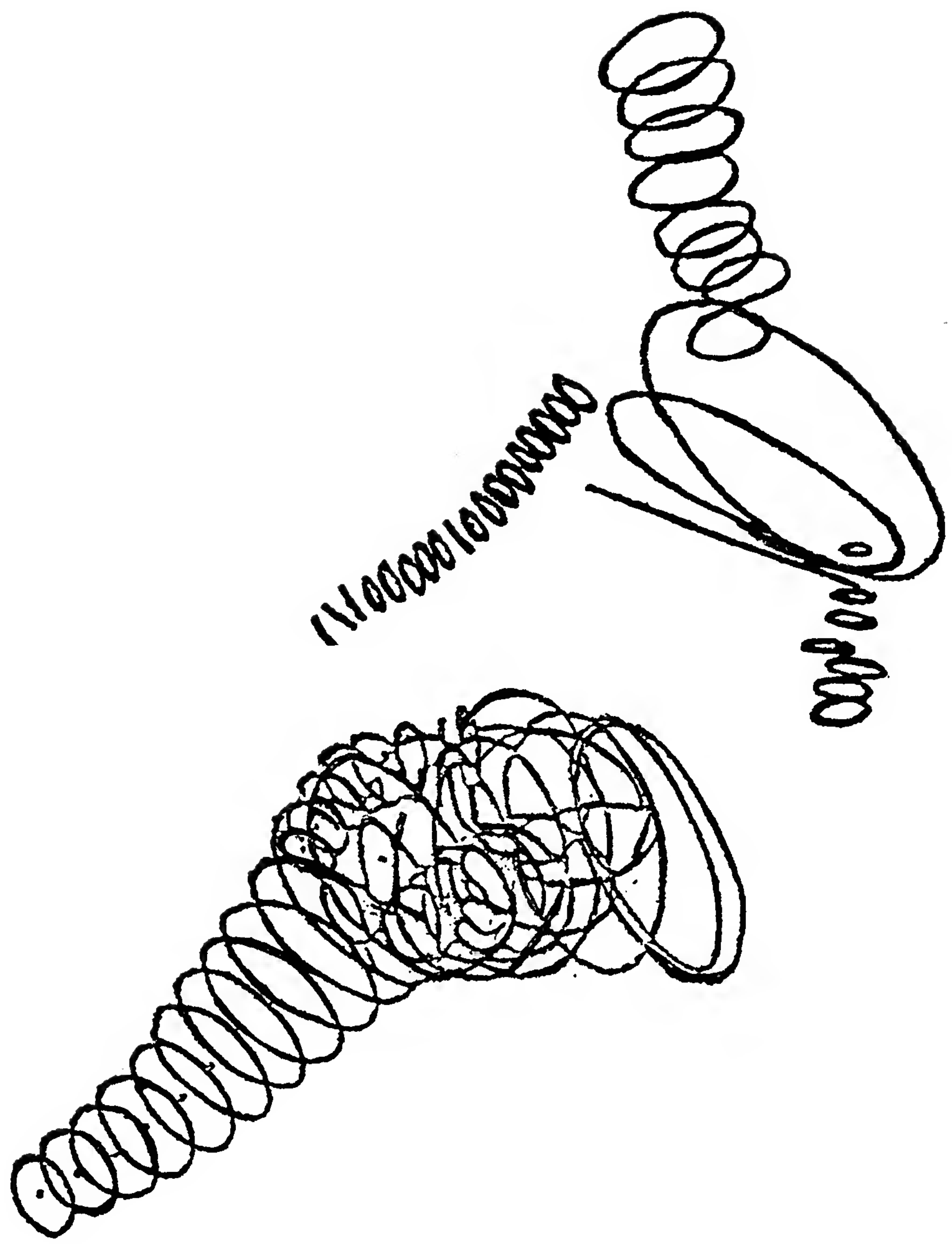
After a matrix of points has been obtained, internal points (points within the boundary obtained from the laser lines) are linked into lines by a "maximal minimal distance" method, which seems to rely on the lines of intensity points. The rough lines are then segmented and approximated by second-degree polynomials (figure 2.17A). Grouping these lines by parallelism yields initial cylinder candidates (figure 2.17B). Axes are estimated by plotting the midpoints of segments. Cross sections are fit at the axis point estimates, and a given cylinder is extended as far as possible (figure 2.17C).

His results show that cylinders are often combined when they should have been separated, such as the legs. At other times the cross section fitting and extension have ill-defined beginning and end points. His techniques seem to work best with single cylinders possessing circular cross sections. When there are multiple cylinders, then obscuration or closeness of cylinders can lead to a poor segmentation. Once again, the



A. B.

FIGURE 2.17. Various stages of Barbie doll visual analysis, from Agin (1972).



C.

FIGURE 2.17. Continued.

root of the problem with his approach I think lies with internal points to guide segmentation.

Nevatia [1974] has improved upon Agin's approach by using contour for both axis estimation and cylinder segmentation. Using the same low-level system as Agin, his analysis departs from Agin's after grouping internal lines by parallelism. This grouping provides a preliminary segmentation, which may later be modified through examination of contour. For each group, a boundary is constructed from the ends of the internal lines. An initial axis estimate is provided by taking the midpoints of the internal segments, and is corrected by constructing cross sections normal to the axis at the midpoints and computing their intersections with the boundary. This process is iterated until it converges to a reasonably stable axis estimate.

Once an axis is found, it is extended a little in each direction and corrected as above. A radical change in radius of cross section is grounds for segmentation. When a single cylinder is thus completed, rough shape descriptors such as axis length and ratios of length to average width of cross sections are computed. Polynomial descriptions are given to axis shape and cross section function: straight or parabolic for the axis, and constant or linear for the cross section function. The joints of the various cylinders are finally computed. Matching against models is conducted by examining the number and structure of single cylinders, and by examining the correspondence between rough shape descriptors.

2.4 Description

Each of the domains line, region, and volume poses unique problems in assigning prototypes to its objects and in bringing about the appropriate modifications. The subsequent three sections treat these domains separately. The discussion is carried out in the context of generalized cylinders, where all three domains play a role.

A common problem in drawing up qualitative descriptions for each domain is boundary fuzziness between categories. Whereas the relative differences between categories are clear, such as between broad and narrow widths, the exact boundaries are not. A boundary must be set nonetheless, and any choice leads to certain problems discussed in section 3.5.

2.4.1 Description of Curves

Contour must be represented in a manner that facilitates description. Quantization of curvature is one way of bringing out general trends, and results in segmentation of curves into quantized segments.

Through investigating archeologists's descriptions and in formalizing curve description for computer, I have concluded that five curvature levels are adequate for most purposes. These are:

```
(line curvature curved strongly)
(line curvature curved round)
(line curvature curved gently)
(line curvature straight fairly)
(line curvature straight very)
```

Note the similarity to the modifier form in section 2.1:

```
(prototype modifier-type modifier submodifier)
```

Here *line* is not so much a prototype as it is a domain indicator.

The two modifier terms are ***curved*** and ***straight***. The standard curvature for curved lines is defined as ***round***; whether a line is strongly or gently curved is measured relative to it. A straight line may be very or fairly straight. Problems in assigning curvature level are discussed in section 3.3.

LINES ARE SEGMENTED AT INFLECTION POINTS

Complex lines are segmented at inflection points into pieces that are assigned qualitative curvature labels. Unfortunately minor line fluctuations give rise to inflection points that could cause segmentation into too many parts, and so it is necessary to smooth the line to average them out. Size might identify such fluctuations because they normally yield very short segments. Some way of summarizing systematic irregularities is also needed, such as saw-toothed, ribbed, or just jagged, but I have not pursued this topic.

Even lines without inflection points may require segmentation, as when curvature varies considerably with length: for example, from strongly curved to very straight. With the type of objects allowed in this thesis, it has been my experience that 2 quantizations suffice to describe such segments.

THE RATE OF TRANSITION BETWEEN CURVATURE LEVELS IS SPECIFIED

Most natural objects vary gradually in curvature; the rate of change

of curvature is as small as possible along the contour. To complete the description, therefore, a transition from one level of curvature to another and the quickness of this transition should be specified. The transitions I have chosen are

```
(becoming abruptly very)
(becoming abruptly)
(becoming)                 :some standard transition
(becoming gradually)
```

When the transition between segments is (becoming abruptly), the term ***corner*** is used. If the transition is (becoming abruptly very) and the two segments are reasonably straight, i.e.:

```
    (straight very)
    (straight fairly)
or  (curved gently)
```

the term ***angular*** is applied. That is to say, an angle is a very sharp transition between two lines that are fairly straight. If the transition was sharp but the lines curved, then the junction would more properly be labeled ***cusp***.

OTHER WORK ON CURVES

Gardin [1967] proposes a differentiation of curvature into five levels (figure 2.18): a-strongly convex, b-slightly convex, c-straight, d-slightly concave, and e-strongly concave. Strictly speaking, convexity and concavity take more into account than just curvature, so that the five levels reduce to three: strongly curved, slightly curved, and straight.

Gabriel [1973] approximates curves with circular arcs. The

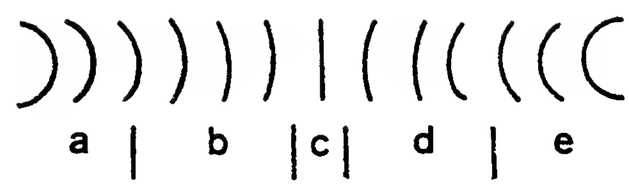


FIGURE 2.18. Differentiation of curvature into distinct levels, taken from Gardin (1967).

psychological objection to this approximation is the sharp discontinuity in curvature between two conjoined circular arcs. For perhaps similar reasons, approximating curved lines with straight line segments also makes people unhappy. Their associative response indicates that curvilinear and rectilinear shapes belong to distinct stimulus domains. "Curves (like poems) lose something in translation" ([Zusne 1970] p.318).

2.4.2 Axis and Cross Section Description

The present section assumes a constant scale change function as in the handles of vases. Scale change is discussed as a separate issue in the next section, since it leads to volume concepts.

AXES ARE ONLY ROUGHLY DESCRIBED

Archeologists do not describe complex axes in great detail; in fact, the more complex the axis, the more approximate its description. A small repertoire of highly approximate prototypes, such as bow, hook, arch, reflex, and stirrup (figure 2.19) is applied practically without modification to handle axes. This repertoire can be represented by the curve quantization of the last section.

Often a general term such as *loop*, which is any axis attached at both ends, suffices as a description. There is great leeway in axis shape because: (1) handles serve a manipulative function, and (2) the ability of handles to serve this function is not strongly reliant on axis shape. Exact shape is therefore relatively unimportant for recognition purposes. The only information normally required about handles is their number, their position, and a rough description such as loop.

Greater approximation with complication can be rationalized as resulting from a lack of constraint among features. Individual features also cease to have any constraint on the name of a vase. These features, when not isolated, may receive a gross characterization such as ***wrinkled***.

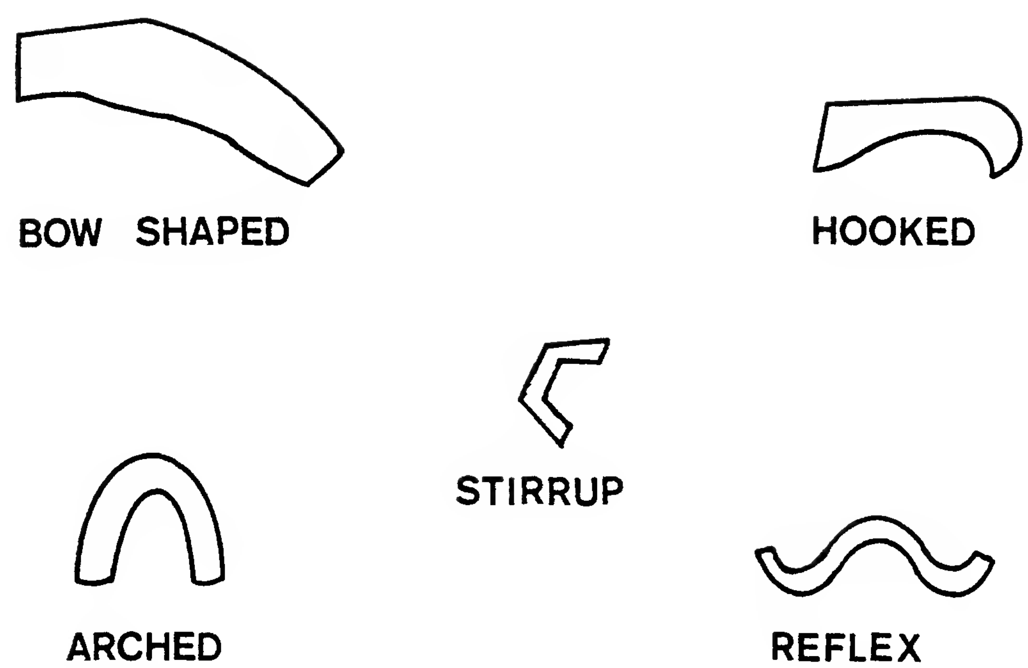


FIGURE 2.19. Some common handle axis prototypes.

SIMPLE GEOMETRIC SHAPES SERVE AS REGION PROTOTYPES

Common regular shapes seem to make the best prototypes, such as rectangle, square, parallelogram, circle and ellipse. The first three are suited towards the polyhedral domain, while the latter are the most generally useful prototypes for curved objects.

I have addressed problems of prototype assignment in the polyhedral domain in an earlier work [Hollerbach 1972b]. Two types of modifiers to regular planar shapes were proposed: indentations and protrusions. Intuitively speaking, a shape can be rigidly modified by cutting something out of it (indentation) or by sticking something onto it (protrusion). Interesting problems result from a fuzzy region between indentations and protrusions; a rectangle with protrusions may with a slight change in protrusion dimension appear to be a square with an indentation (figure 2.20). Some precise results were obtained about this fuzzy region, and are presented in section 4.

PEOPLE USE THE SAME PROTOTYPES

The experiments of Rosch [1973] indicate that such basic forms as listed above serve as prototypes across all races and societies of people. Her subjects were members of the primitive Dani tribe of Indonesian New Guinea. They do not possess terms in their language for simple geometric forms, and do not appear to have "unspoken" concepts for them. The experiments involved selection of the most typical member from a set of similar shapes, such as may be obtained by modifying a square (figure

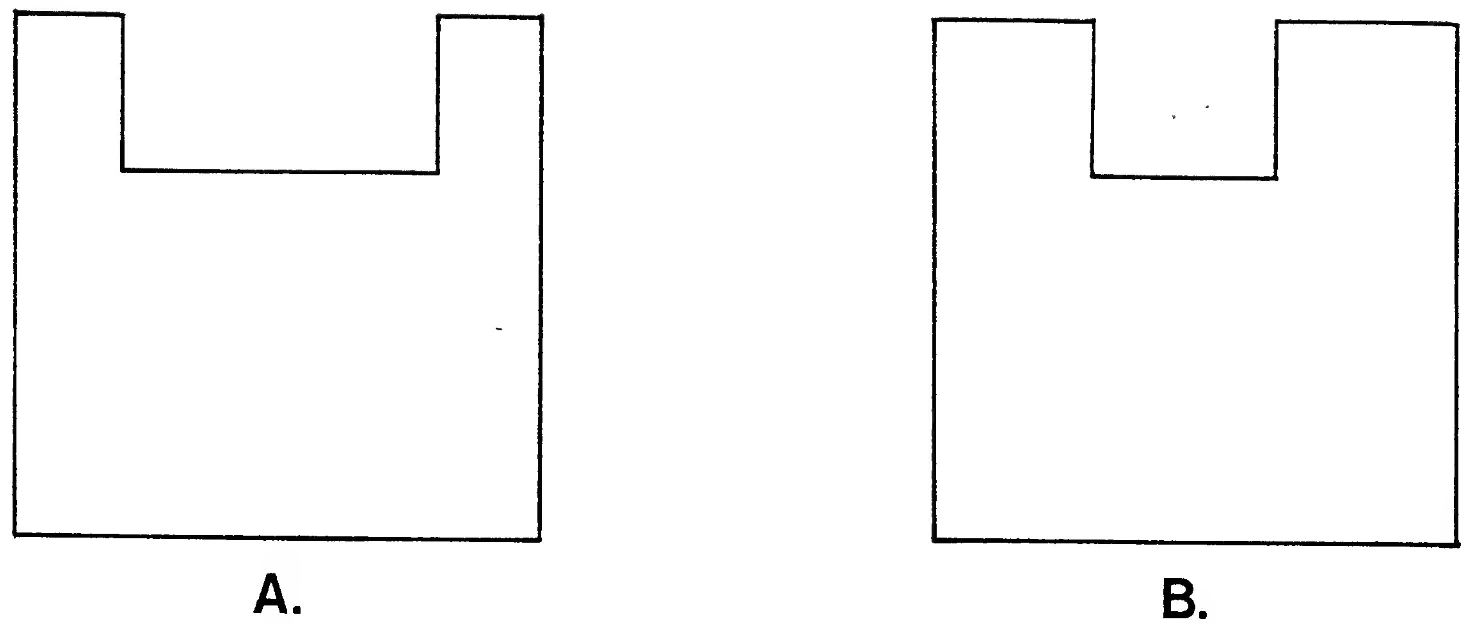


FIGURE 2.20. Region A is most often judged by people as a rectangle with protrusions, while B is considered a square with indentation.

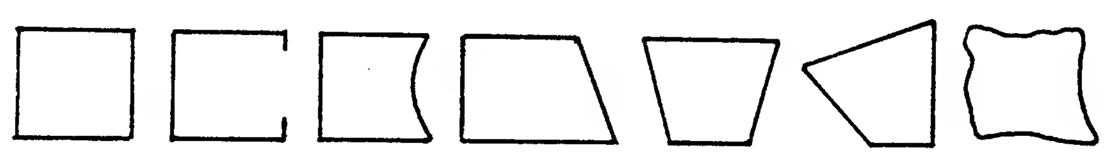


FIGURE 2.21. Basic square with six modifications, taken from Rosch (1973).

2.21).

Her results showed that the simple geometric figures were almost always chosen as most typical--an argument for the existence of natural prototypes. Control experiments were run to ensure that the Dani do not have a preexisting bias towards grouping 2-dimensional figures into form classes. That they do not results perhaps from their living in an "uncarpentered world" that contains only irregular 3-dimensional shapes and no 2-dimensional objects or figures. Descriptive economy explains these results: those shapes with simple descriptions more readily serve as common denominators between diverse shapes than more complicated ones.

2.4.2.1 Other Work on Region Description

Gardin [1972] has suggested some primitive cross section shapes and decorations for handles (figures 2.22 and 2.23), based on a survey of use by archeologists. His suggestions can be interpreted in terms of modifications to elliptical and circular cross sections. Cross sections 12 and 13 (figure 2.22) can be interpreted as modifications to a standard ellipse, obtained by altering the ratio of major to minor axis and the boundary shape. Cross section 15 can be considered as two overlapped circular regions. The decorations suggested by Gardin are actually of the two modifier types indentations (2p, 2q, 3p, 3q) and protrusions (1q, maybe 5p). Notch and finger depression are two different types of indentations.

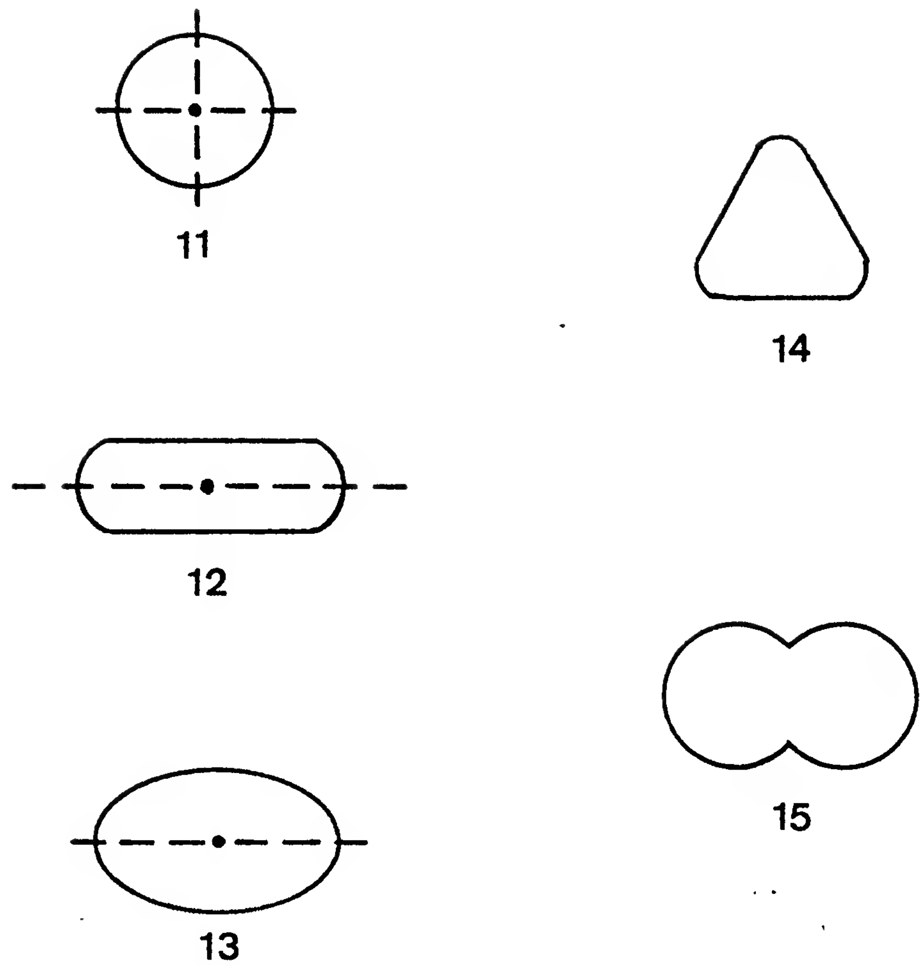


FIGURE 2.22. Handle cross section proposed in Gardin (1972):

11 circular
12 flat
13 oval
14 triangular
15 geminate

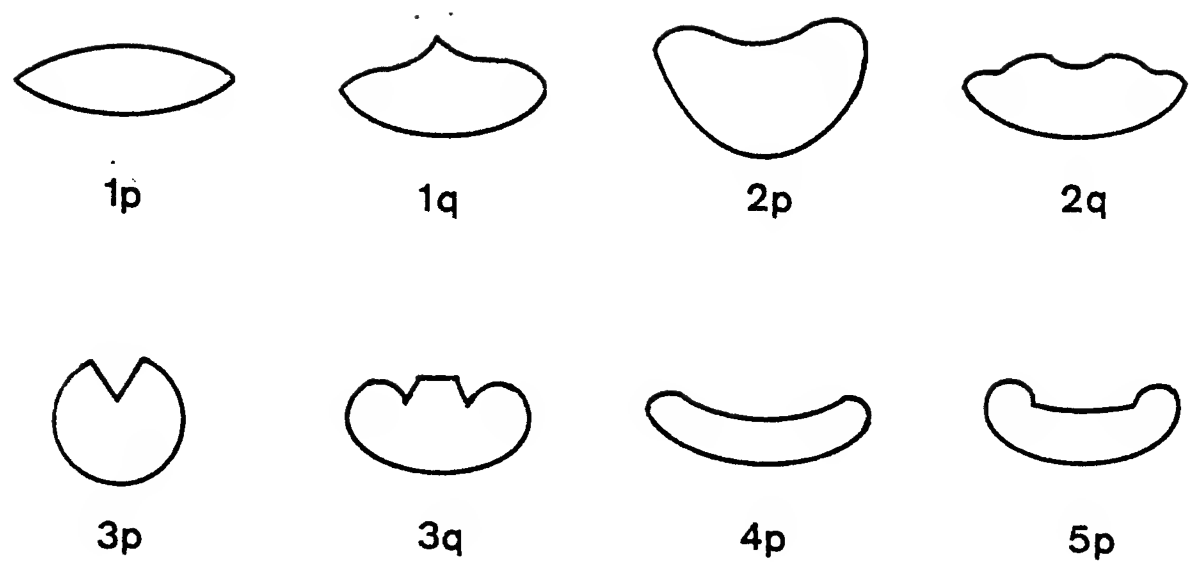


FIGURE 2.23. Handle cross section decorations, taken from <u>Gardin</u> (1972):

1p two arêtes situated laterally
1q an arête situated centrally
2p an impression of the finger
2q multiple impressions of the finger
3p a notch
3q multiple notches
4p arched section
5p section with a flat strip

REGION PARAMETERIZATION

Much work has been done on region parameterization. Maruyama [1972] lists some quantitative measures and their proposed interpretation:

1. jaggedness, P^2/A, where P is the perimeter length and A is area;
2. degree of skewness, which corresponds to the third moment of area;
3. degree of elongation, corresponding to the fourth moment of area.

Zusne [1970] reports that the second moment of area or of perimeter about the x or y axis has correlated well with major axis estimates. Krakauer [1971] uses an eccentricity measure to describe the shape of his regions.

The problem with these parameters is that they do not pin down shape exactly enough. Wildly different shapes may give the same parameter value; for example, a deeply convoluted figure could give the same jaggedness value as a very thin rectangle or ellipse [Attneave 1956]. When moments of area are computed, moreover, the sheer size of the area enclosed obscures small perimetric features [Maruyama 1972]. Local features may sometimes be unimportant but at other times represent a significant portion of the description.

THE MINOR DETAIL BUG STRIKES AGAIN

Two other approaches have the opposite problem: they are too sensitive to local features. Guzman [1970] simply models a region as a concatenation of segments that form the boundary. This involves segmentation of the boundary into distinct line segments and description of

each line by a chain-coding scheme. Objects are modeled as a collection of such regions. There are two serious problems with this approach. (1) It is difficult to compare regions that differ only in minor detail, since such detail can induce widely different segmentations or line descriptions. He needs multiple templates to represent possible appearances of a model; note, for example, the collection of templates to represent a hat in figure 2.24. (2) Perspective deformation can change the apparent shape of the boundary.

The second approach is the medial axis transform [Blum 1964]. A skeleton is generated for a region by connecting the centers of discs that satisfy two conditions: (1) the disc is the largest possible one centered at a particular point while still being within the boundary; and (2) the disc is not completely contained by some other such disc. Although this transform has been extended to 3 dimensions, the objections to the two and three dimensional versions are the same. Agin [1972] has nicely summarized them. He notes that small changes in contour bring about great changes in the transform of regions, for example, transforms of a rectangle with and without notch (figure 2.25). Finally, the descriptions are highly unintuitive and hard to use.

The minor detail problem is thus seen to wreak havoc with both Guzman's and Blum's approaches. The difference between a rectangle and a rectangle with a notch is the notch.

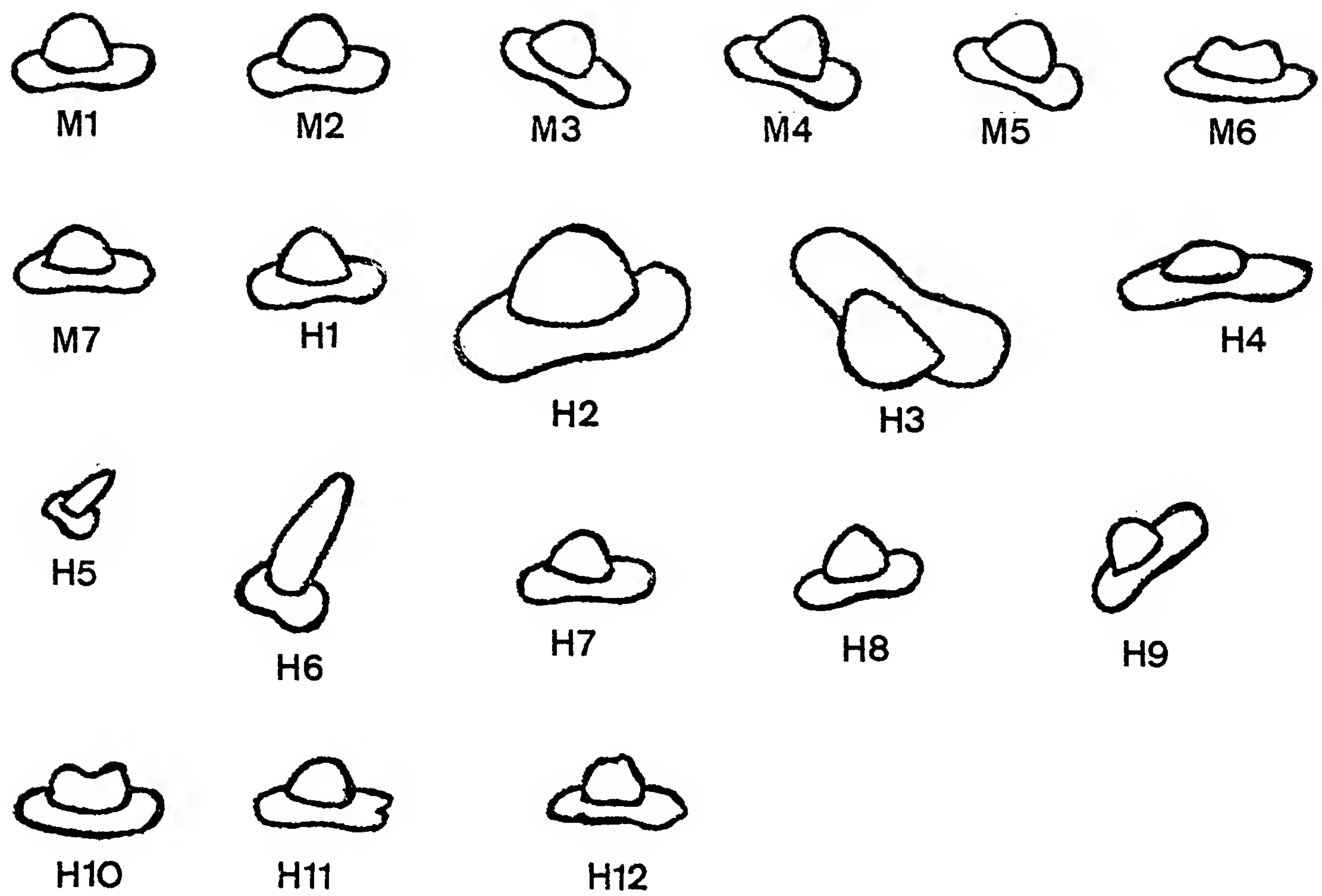


FIGURE 2.24. Some templates to represent a hat, taken from Guzman (1970).

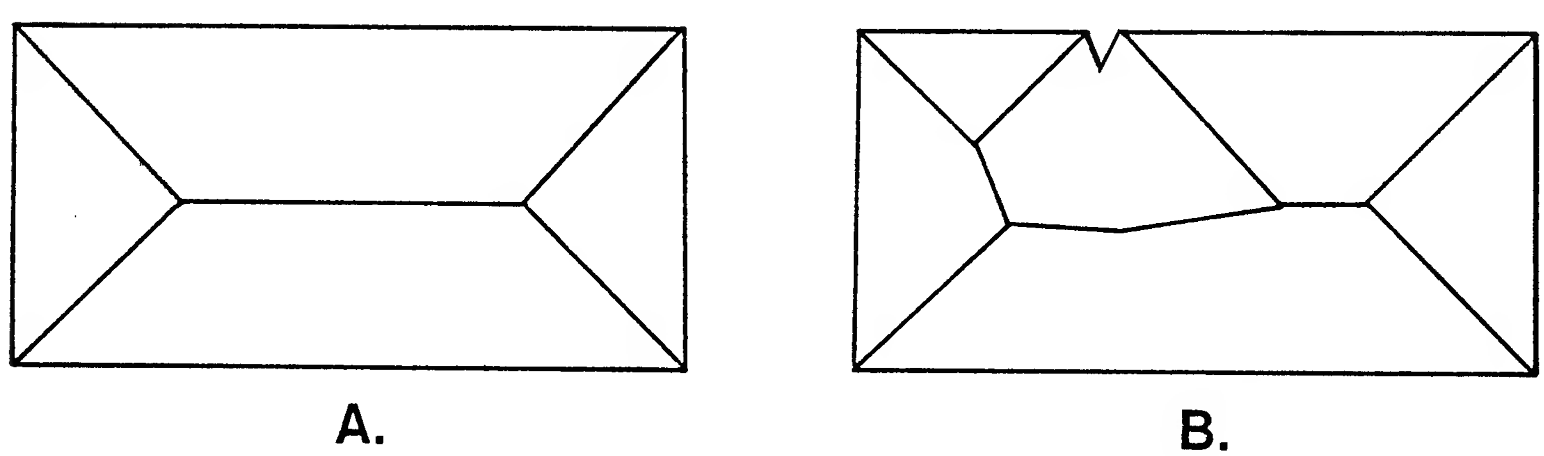


FIGURE 2.25. Blum transform of a rectangle (A) and of a rectangle with a notch (B), from Agin (1972).

2.4.3 Cylinder Parts

This section deals with scale change for the special case of circular cross section and straight axis. This case is the most common and important one, and serves as a default condition on cylinders. If a cylinder departs in minor ways from a straight axis or circular cross section, it can be described in default terms along with additional modifiers. Archeologists do a little of this, speaking of a body as flattened when the cross section is elliptical. Otherwise, if the axis and cross section are complicated, it is better to describe them explicitly than to give the type of description presented below.

THE SOLID PROTOTYPES

To obtain a broad overview of what the scale change function is doing and to place irregularities of outline in perspective, a set of prototypes and modifiers must again be devised. The mathematically simplest forms of scale change are constant, linear, and quadratic functions of the axis. When coupled with a straight axis and circular cross section, they yield the familiar cylinder, cone, ellipsoid, paraboloid, and hyperboloid. The distinctions among ellipsoid, paraboloid, and hyperboloid, however, are too specialized to be of use for qualitative description.

Archeologists commonly use *cylinder cone* and *ovoid* prototypes. *Ovoid* corresponds to an ellipsoid deformed to leave one end bulkier than the other (figure 2.26). The concept of *bowl* a shape whose top is wider than the bottom and whose height is considerably less than the width, is common

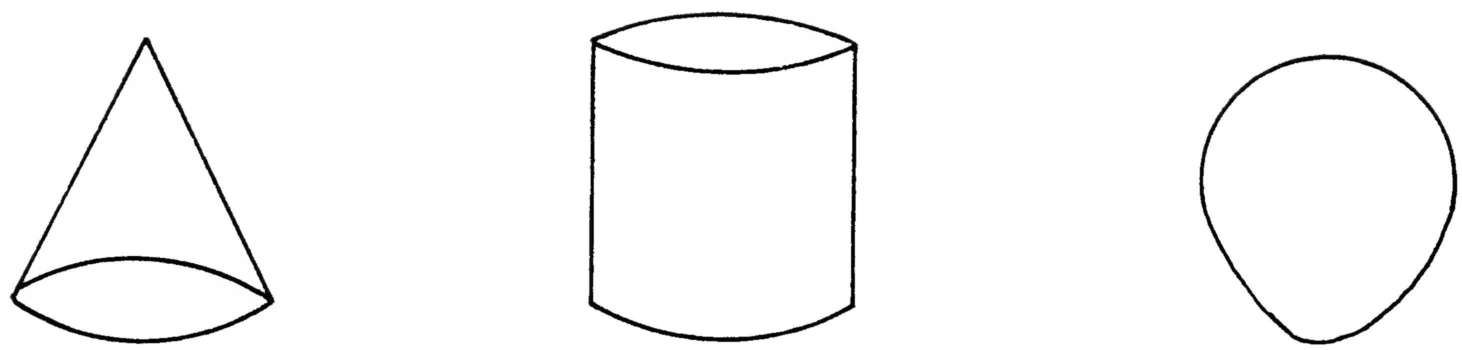


FIGURE 2.26. Prototypes cone, cylinder, and ovoid.

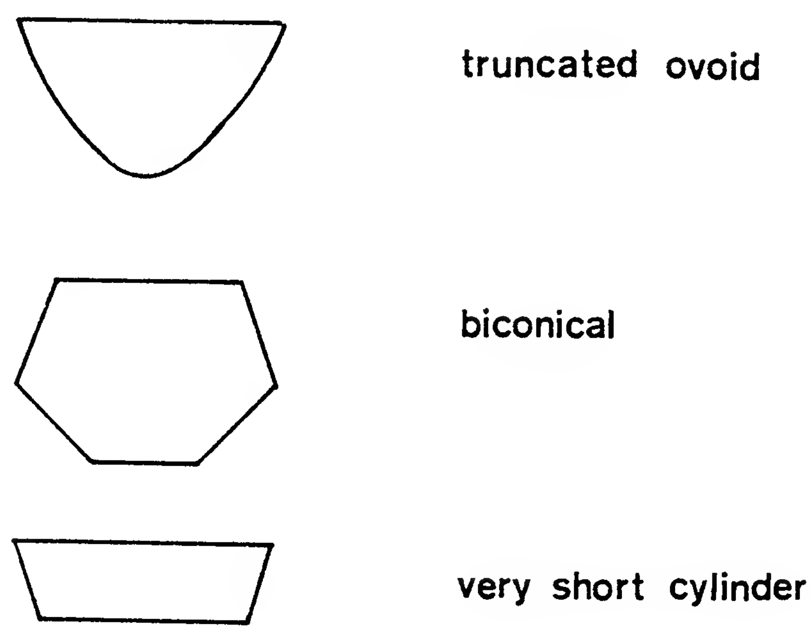


FIGURE 2.27. Some common bowl shapes can be described in terms of the other prototypes.

enough to merit its own prototype, even though common bowl shapes can be described by the first three prototypes (figure 2.27). Moreover, the distinction between open vases (bowl) and closed vases (ovoid) is fundamental in archeology.

Less common prototypical shapes are spherical, hemispherical, biconical, piriform (pear shaped), and bell shaped. Because archeologists employ these terms, and because they are common in everyday language, these shapes have also been incorporated into the description programs. Some are actually considered modifications of other prototypes; for example, sphere is a special case of ovoid. The exact character of the modifiers that suggest these lesser prototypes is given in section 3.3.

MODIFIERS ARE ASSOCIATED WITH EACH PROTOTYPE

Some types of modifications are common to all prototypes, some are prototype specific (see table 2.1 at the end of this section). Note once again that the structure of a modification is that proposed earlier. The choice of submodifier term is flexible, and a number of more or less equivalent ones are in use: relatively and fairly, sharply and strongly, gently and mildly.

The modifiers are expressed in general terms to bring out underlying relationships, although the exact terms may differ from prototype to prototype. For example, a tall bowl is usually referred to as a deep bowl, a short ovoid as a squat ovoid, a concave cone as a splaying cone. This is discussed further in section 3.3.

There are interesting problems in shape assignment, as there are for cross sections. At some point cones transform into cylinders, cylinders into ovoids, bowls into cones and into cylinders, etc., under the action of modifiers. An attempt at defining these points is deferred until section 3.2.

RIGID VERSUS PLASTIC MODIFIERS

The modifiers for these prototypes are generally plastic deformations, as opposed to the rigid modifiers indentation and protrusion for cross sections. Plastic deformations are natural for pottery, since the soft clay as the vase is made is readily deformed. For example, a plemochoe (figure 2.28) is a large container for perfume used by ancient Greeks and looks like a flattened sphere, and that is exactly how it is made [Noble 1965]: thrown as a sphere and flattened. Otherwise it is easy to make a vase taller, to transform the point of greatest width from low to high, or to give the contour a slight concavity before the clay has hardened.

The only modifier that is not a plastic deformation is orientation for cone. It is a rigid transformation, a rotation, from the standard position of base low and point high to an inverted position.

Truncation is also a rigid modifier. A hemisphere is a truncated sphere. An ovoid may be truncated at the bottom or top to make way for a wider base or neck. When a height-width modifier is assigned to ovoid, allowance must be made for the amount of truncation.

Truncation modifiers have not been included in the list because they

FIGURE 2.28. A plemochoe looks like a flattened sphere.

are implicit in the size of other parts. Bottom truncation of a vase body is indicated by base width, such as broad base, narrow base, or blunt point. Top truncation is indicated by neck or mouth width. How the parts fit together and the appropriate descriptors for conjunction are discussed next.

Table 2.1

prototype	modifier-type	modifier	submodifier
all	height-width	short	very
			extremely
		tall	very
			extremely
all	convexity	convex	
		concave	
		straight	
all	contour	straight	very
			fairly
		curved	gently
			round
			strongly
		carinated	slightly
			sharply
ovoid, bowl, cylinder	shoulder	yes	
		no	
ovoid, bowl	greatest width	high shoulder	
		low belly	
cone, cylinder	slant	vertical	
		slanted in	low angle
			high angle
		slanted out	low angle
			high angle
cone	orientation	standard	
		inverted	

2.5 How the Pieces Fit Together

Once the individual pieces of an object have been described, they are structured into a complete description by specifying relative size, the place of junction of two pieces, and the junction definition. The simplest junction is that between pieces from a single cylinder. Because these pieces share the same axis, one need only specify a one-dimensional position relation, such as above or below, left or right. Pieces from separate cylinders, however, have complete freedom in how they meet. A more elaborate specification of relative position is then required.

TYPES OF CYLINDER JUNCTIONS

Agin has studied cylinder junction for intersecting axes. He calls the point of intersection a *joint*. If one cylinder may move with respect to the other, the joint is called *articulated*, such as a hinge joint. He categorizes joints according to how many axes converge at a joint and whether the axes meet end to end or end to middle.

The main vase cylinder and its handles do not form joints in Agin's sense because the handle axes do not necessarily meet the main axis at a hypothetical intersection. This more general junction is described in this thesis by fixing the main cylinder and by positioning the handle axis relative to it. Positioning a handle involves specifying location (the place on the main cylinder to which the handle is attached) and orientation (the attitude of the handle axis relative to the main cylinder axis).

LOCATION

Since the main cylinder is rotationally symmetric about a vertical axis, a vertical modifier suffices to specify handle location. Since the foot, body, neck and lip of a vase are arranged vertically, location is conveniently specified by referring to them.

In the simplest case, naming the subpart specifies the location, such as neck handles. Finer localization is provided by adding one of the submodifiers ***high***, ***low***, or ***halfway-up***. An ovoid subpart once again has its own special terminology: the halfway point is replaced by the point of widest diameter, the high portion is called the shoulder, and the low portion is called the belly.

The attachment of handles near extremeties of a subpart can be indicated by adding the subsubmodifier ***very*** to high or low. Archeologists describe the situation slightly differently if there is another subpart near the extremity. They say "high on subpart1 near subpart2"; for example, high on the neck near the lip.

The ends of a handle do not necessarily lie on the same subpart. When this occurs the location of each end is given: lip to shoulder, lip to widest diameter, etc. Vertical handles (see below) tend to need such a description.

ORIENTATION

The attitude of the handle axis relative to the vertical main cylinder axis is also specified. When the ends of the handle axis lie on a

horizontal line, the handle is called *horizontal*. When they lie on a vertical line, the handle is called *vertical*. Although other orientations are conceivable, they are not normally encountered in pottery.

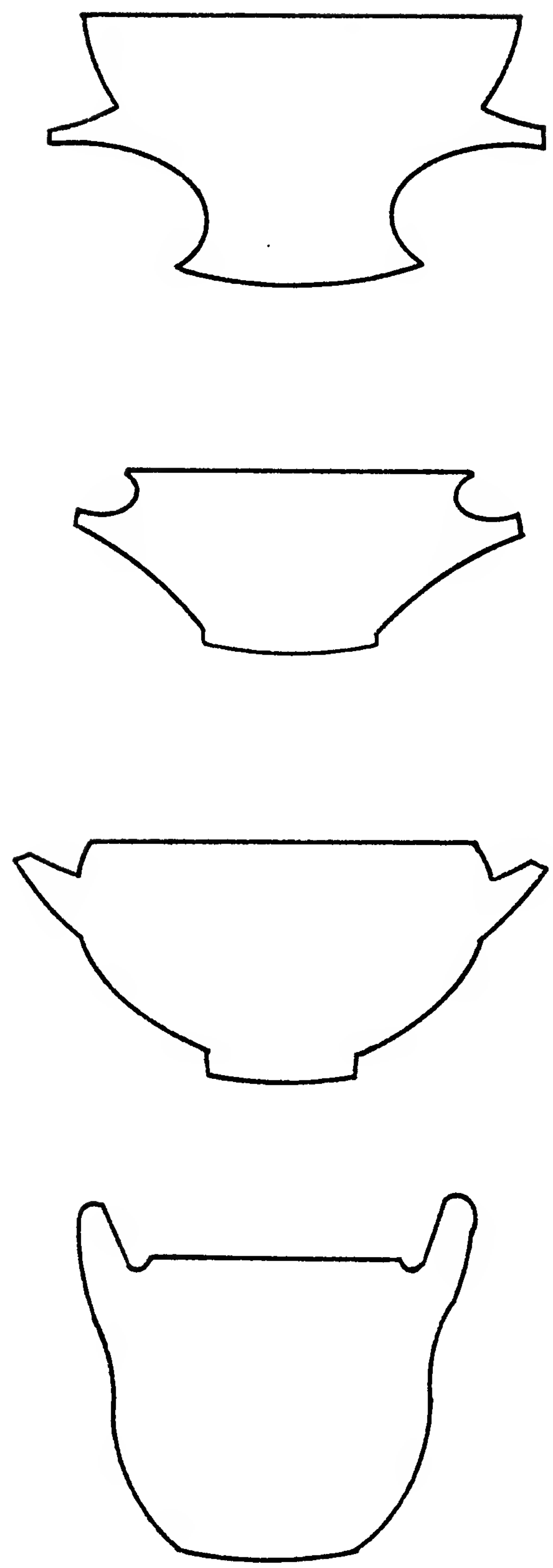
Vertical handles are seldom slanted; that is to say, the main cylinder axis usually lies in the plane of the handle axis. Horizontal handles, on the other hand, are often slanted with respect to a horizontal plane through the handle axis ends. When the handle is slanted below the horizontal plane, the handle is said to slant downwards; when above, it is said to slant upwards. Upward slanting handles are more precisely described by a three-level quantization (figure 2.29): low angle, high angle, or upright angle.

The horizontal or vertical orientation of handles is functional: horizontal handles allow a vase to be carried, vertical ones are good for pouring. Any other orientation would serve neither purpose as well.

ARTICULATION

The junction of two pieces such as body and neck may be sharply defined and angular, or it may be ill-defined as one piece gradually melts into the other. Archeologists describe the junction by the word *articulation*. An articulated junction has two sharply offset pieces, an unarticulated one has a continuous curve between them (figure 2.30). As mentioned earlier, the word *articulation* also describes a movable joint. In this thesis, it is used only in the archeological sense.

There is a real-world basis for the distinction, deriving from how the



A.

B.

C.

D.

FIGURE 2.29. These horizontal handles slant downwards (A.) and rise upwards (B., C., D.). They rise at a low angle (B.), at a high angle (C.), or upright position (D.).

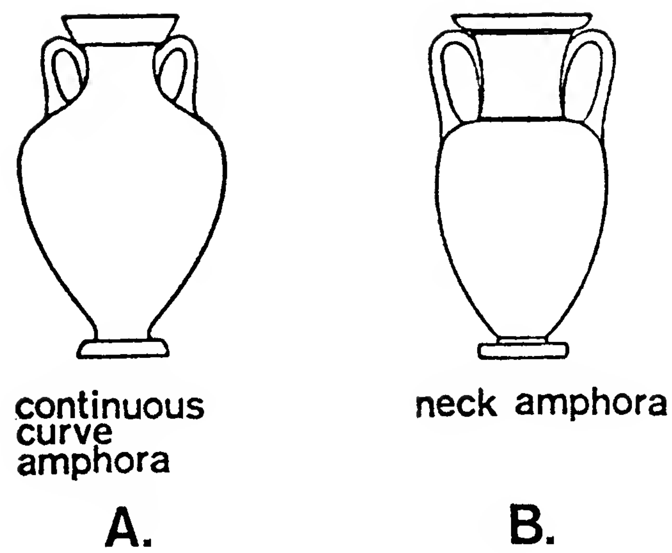


FIGURE 2.30. Continuous curve amphora A and neck amphora B.

vase is made. When a vase is thrown as one piece, the contour tends to vary gradually as one piece flows into the next. When thrown as separate pieces and joined, the junction is much more angular and well-defined. Articulated vases are often factory produced, made in assembly-line fashion [Noble 1965]. The junction between handles and main cylinder is almost always articulated, since they are constructed separately and joined.

RELATIVE SIZE

In describing the relative size of pieces as in describing position, one is chosen as the standard against which to compare the others. Once again, the body of the vase is the standard because of its greater size. An isolated subpart such as the vase body has no standard against which it can be measured, and so a dimensionless quantity like height-width ratio is appropriate to describe its size. The heights and widths of the other subparts are described relative to the height and width of the body.

The height modifier terms are ***high*** and ***low***; the width terms are ***broad*** and ***narrow***. These modifiers may be refined by adding the term ***very***, which leads to a 4-level quantization for each modifier type.

Comparing handle size to body size is made difficult because of curved handle axes, which leave no clear height and width dimensions. Archeologists therefore describe handle volume instead of handle height and width, and apply the qualitative terms ***large*** and ***small***. The terms large and small take curvature into account, and are therefore preferable to ***long*** and ***short***, which refer to a straight-line measure from one end of the axis

to the other.

It thus appears that archeologists give more precise meanings to size terms that are often synonymous in everyday speech. Tall and short refer to a height-width ratio, high and low to a height measurement only, and large and small to volume.

2.6 Flat or Round Shapes

Flat shapes and round shapes do not make very good generalized cylinders. Flat shapes like disks have extremely short axes, which leave scarcely any contour to describe. Flat shapes are encountered in pottery as lips and low feet. Of the limited descriptors one can assign to such shapes (see section 3.4), width is the most predominant.

Spheres make poor generalized cylinders, because, as Agin has remarked, it is difficult to select an axis as the predominant orientation. Such rounded shapes are found in pottery as lugs. A lug is a form of handle that is grasped by pinching or that is pierced for suspension purposes. The grip angle or the pierce is normally horizontal or vertical. Some lug profiles are given in figure 2.31. Rough prototypes may be assigned to lugs such as the bowl shaped lug in figure 2.32A or the horned one in figure 2.32B.

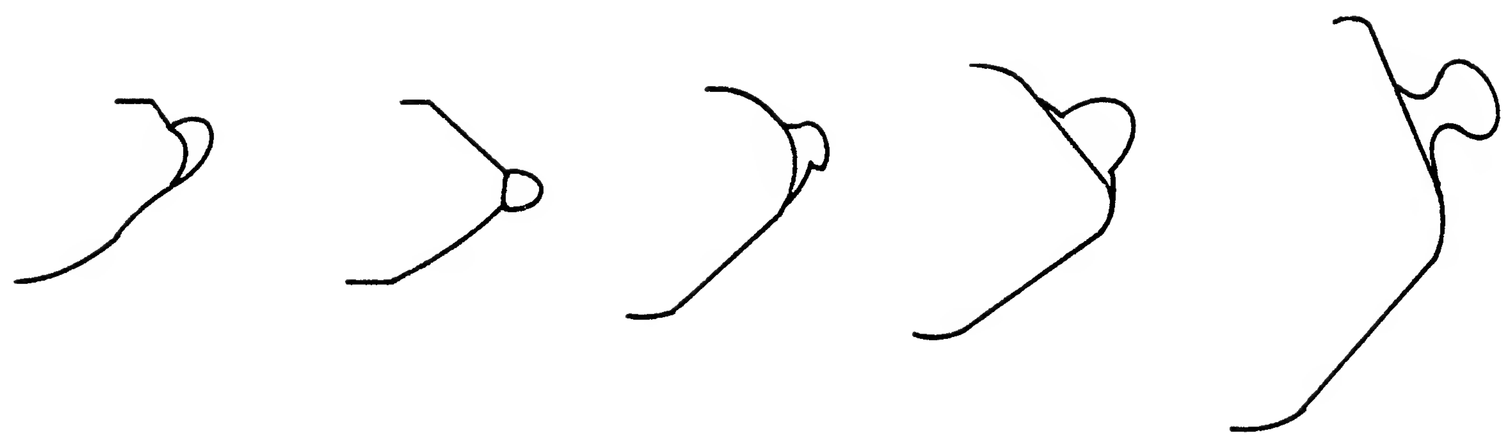


FIGURE 2.31. Profiles of assorted lugs on the shoulders of bowls, taken from Warren (1969).

FIGURE 2.32. Two lugs showing similarity to a bowl shape (A.) and to a horned shape (B.).

2.7 Relation to Psychological Work

The notion of prototype finds scattered mention throughout the psychological literature. The concept of ***schema*** is due chiefly to Bartlett [1932], and has startling analogies to Minsky's [1974] frame systems. Bartlett's schema provide "an appropriate frame" to the material in question. It provides the first general impression, and sorts the general tendency from the details. Elaboration of detail follows only after the setting has been laid. Bartlett also noted the effect of schema choice on what is perceived, namely, that there are associated with a schema conventional representations which determine the interpretation of detail. This is like the preexisting slots or default assignments of a frame system.

Woodworth [1938] spoke of ***schema with correction***, where his use of the word ***schema*** is more precise and restricted than is Bartlett's, and is close to the present formulation of prototype. After considering a number of experiments on memory of form, Woodworth concluded that a geometric configuration is usually remembered by assigning it a schema, a simple geometric form, which is then corrected. A figure might be described as "a square with a nick on one side".

A brief mention of this type of description appeared very early in Kuhlmann [1906], where subjects were observed to remember shapes as altered familiar forms. Neither Kuhlmann nor Woodworth, however, developed the idea. In later editions of Woodworth's book the terms schema and correction disappear entirely in favor of Gestaltist concepts.

A later mention of schema with correction appears in Hebb [1949]. He notes that subjects perceive a pattern, first, as a familiar one, and then with something missing or something added: for example, "a triangle with the top cut off" or "a square with a crooked bottom."

Whereas Woodworth's mention of schema was drowned out by Gestaltism, Hebb's mention of it was buried by the impact of information theory on perception. Forms were reduced to numbers that represented their degree of complexity, and from these numbers were magically supposed to emerge theories of perception. Not only psychology was infected with this approach, but also machine vision in the form of pattern recognition. The inability of information theory to account for the complicated processes of vision, however, gradually became apparent.

Towards the end of the application of information theory to perception appeared another mention of schema and correction, this time in the work of Gombrich [1965]. He develops the idea extensively in the domain of visual art. A schema to him represents the first, approximate, loose category which is gradually tightened to fit the form it is to reproduce. It is not the product of a process of abstraction, of a tendency to simplify, as information theorists would have it. His schema are preexisting things or concepts, so that perception is primarily the modification of an anticipation.

PROTOTYPES IN CURRENT PSYCHOLOGY

If the frequency of use of the term *prototype* in current psychological

literature is any guide, this concept's day has finally come. As discussed earlier, Rosch's interpretation of the meaning of prototype is similar to my own. Posner [1968], however, uses the term in an information theoretic sense that is opposed to the spirit of my usage.

Posner's work is an elaboration of some early work done by Attneave [1957]. Attneave was one of the strongest proponents of the information theoretic approach towards perception (see Attneave [1954]), and his prototypes, or schemata as he calls them, are creatures of this approach. A prototype is supposedly that pattern which has the most in common with the other patterns of a group, i.e., that pattern for which the sum of variations between it and the other patterns is the least. However, this prototype is not fixed, it has no structure, and it varies with membership in the group. It is not clear what the description of the prototype is, or exactly why it is a prototype. All we have is an obscure sum of variations.

Prototype as used in this thesis is a preexisting form, fixed but modifiable. Descriptions may vary, but prototypes do not. What Posner and Attneave evidently have in mind is the most typical member, where "most typical" is determined by some statistical measure. Posner shows by statistics that subjects learn or remember his prototype easier than other patterns of the group, and claims that this shows information common to individual instances is abstracted and stored in some form. He has not shown how this is actually done, which is the really important question.

CHAPTER 3 -- POTTERY

This section presents the description methodology as applied to vases. A program has been written to describe and identify vases from their outlines. The program consists of 4 stages:

1. Segmentation into foot, body, and neck or lip.
2. Prototype selection for parts.
3. Modifier assignment to prototypes.
4. Function and name assignment.

The subsequent sections detail these stages.

DESCRIPTIVE TERM BOUNDARIES

Two basic difficulties have been encountered in this work. One is to give precise meanings to qualitative or fuzzy terms by setting a quantitative boundary between descriptors of the same type, such as between broad and narrow. The other difficulty is to get around these boundary definitions when there is a borderline case. Narrow-necked vases, for example, normally receive different classifications than broad-necked vases. When a neck width is near the border line of narrow and broad, it becomes somewhat arbitrary which assignment it receives, since the border line itself is somewhat arbitrary. One must be prepared to treat the neck width either way, and to abandon one width assignment for the other when mitigating circumstances arise.

Though troublesome, most of the time the boundary problem will not arise. The qualitative distinction between terms is usually clear and provides a useful basis for making decisions. Most situations will not lie near the boundary, but at a comfortable qualitative distance from it.

The rough location of a boundary may be fairly important, although exact positioning is not. A boundary may violate real world constraints that favor an approximate location for distinguishing vase forms. The distinction between narrow and broad necks, for example, is based on the properties of liquids versus solids. Narrow-necked vases make for greater ease of pouring and for transportation without spillage. Broad-necked vases are more suited for entering or removing solid material.

VASE CATEGORY BOUNDARIES

Setting boundaries for functions and names is more difficult than setting descriptive term boundaries. One reason is less precise definitions and usage. A dictionary definition of jar, for example, is an earthenware container having wide mouth and often no neck. Yet some vases having this description are not called jars, while some jars deviate from this definition by having narrow mouths.

Another reason is that several descriptive dimensions are involved in a vase category name. Because of limited evidence it is hard to decide when a particular dimension has exceeded the limits for that category and transformed the vase into a different type. For example, the dividing line between kylix and skyphos, two Greek drinking cups essentially

distinguished by depth of the bowl, is unclear. In drawing up category boundaries, I have as a result had to rely heavily on intuition.

BOUNDARIES AND ARCHEOLOGICAL USAGE

Insofar as possible, archeological usage was observed in setting boundaries. The terms were largely derived through study of *Greek Geometric Pottery: A Survey of Ten Local Styles and Their Chronology* by J. N. Coldstream. His descriptions are particularly rich and consistent. These terms were augmented and refined by examining Lacy [1967], Noble [1965], and Warren [1969]. Cook [1960] and Richter and Milne [1973] helped in delineating Greek vase categories.

From this study, I deduced that for the most part archeologists use similar terms in a reasonably consistent structure: hierarchical, based on selection and modification of prototypes. This consistency has made it possible for me to come up with a set of terms, precisely defined, that correspond well with archeological descriptions and everyday usage. The vase descriptions derived by my program are consequently natural sounding, and are comparable to what an archeologist would give.

Though archeological descriptions can be formalized, archeologists as a whole appear unaware that they are using a consistent structure or that they are applying descriptive terms fairly precisely (an exception is Gardin [1972]). This implicit formality made it difficult for me to pinpoint a boundary: sometimes contrasting terms as seen in different vase descriptions overlapped; sometimes all examples of a particular set of

contrasting terms that I could find lay too far apart to pinpoint a boundary. Adding to this difficulty is that archeologists more often give comparative than absolute descriptions. They more commonly describe a neck as broader than some other neck than they describe a neck as broad or narrow. Thus I have often had to set a boundary by analogy with similar but more exactly related terms, or by substituting personal impressions. Lack of explicitness in definitions, I might add, is causing archeologists difficulties in recent attempts to computerize vase holdings by museums Whallon [1972].

The program does not segment and describe handles, although handles are important in function and name assignment. This involves detecting handles in all sorts of positions--partially obscured, within the boundary of the main vase cylinder, etc.--and I was not prepared to deal with this generality of position. Handle descriptions as discussed in section 2 are externally provided to the function and name assigner, although the program itself provides the main cylinder description. Finally, no provision has been made for spouts and lids.

3.1 Segmentation

The present section is concerned with segmenting the main vase cylinder into three parts: a foot assembly, a body, and a neck assembly. Further segmentation of the foot and neck assemblies is discussed in section 3.4. A pedestal foot is broken into base and stem; a neck assembly may be split into neck and lip.

Outlines of vases are entered to my program as lists of points. Since a vase is symmetrical about its axis, only the half profile need be entered. For the amphora of figure 1.2, duplicated without handles in figure 3.1A, the outline as entered is shown in figure 3.1B. The points were manually computed from the smooth outline. Values were quantized coarsely for convenience of entering these points, although some jaggedness of the point list resulted.

Because the cylinder's axis is vertical and straight, one can speak of width change instead of scale change. To locate regions of large width change, the cylinder axis is first divided into unit intervals. For each interval, the change in width (abbreviated DW in the figure) is computed by differencing the width values at the ends of the interval. The neighboring DWs are then differenced to yield the rate of change of width (abbreviated DDW).

CHOOSING THE SIGNIFICANT WIDTH CHANGES (DWs)

Ignoring the large width changes resulting from the flat top and bottom of the base, the average DW over all the intervals is 1.3. There

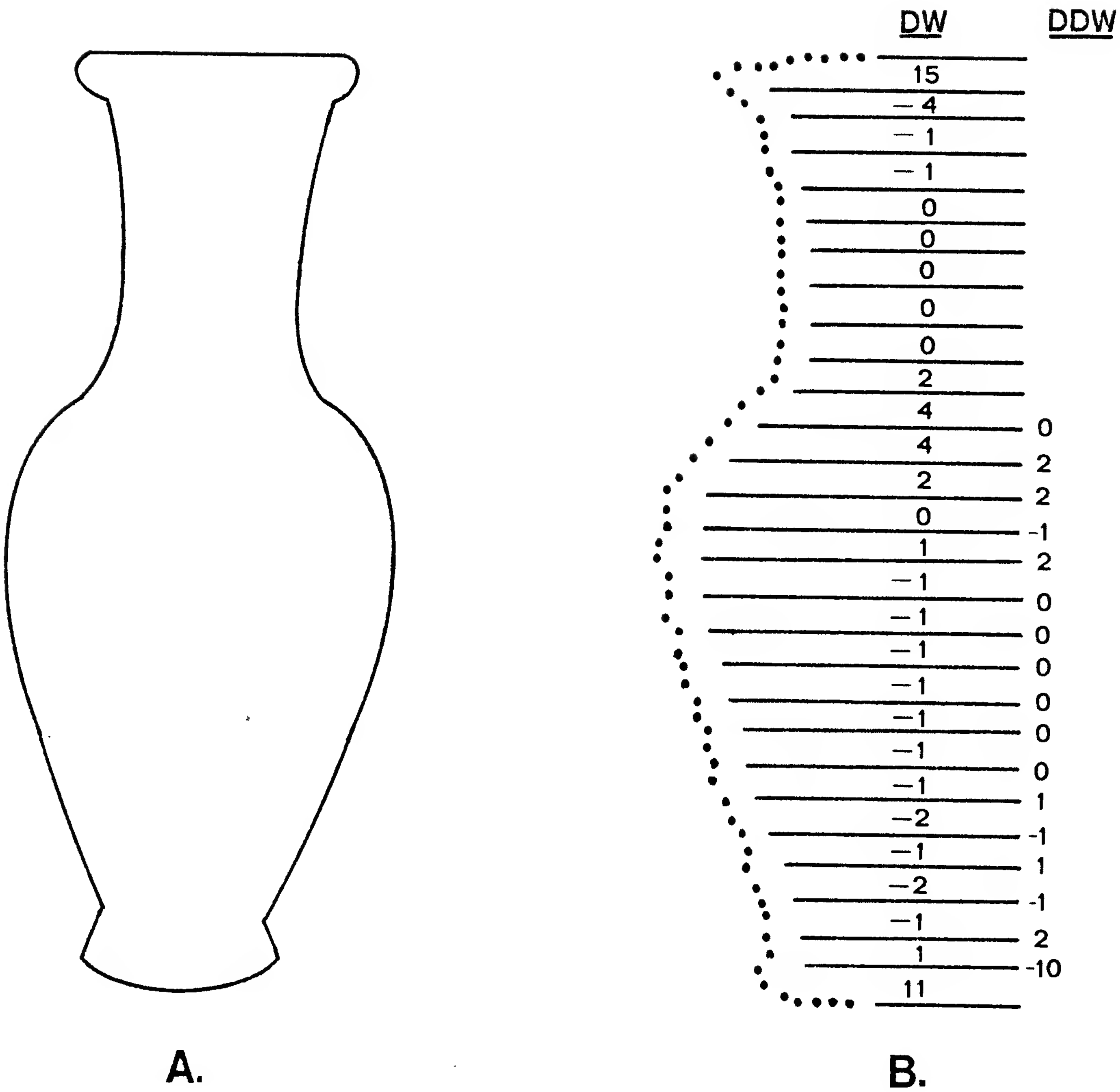

DW
DDW
15
−4
−1
−1
0
0
0
0
0
2
4
0
4
2
2
2
0
−1
1
2
−1
0
−1
0
−1
0
−1
0
−1
0
−1
0
−1
1
−2
−1
−1
1
−2
−1
−1
2
1
−10
11
A.
B.

FIGURE 3.1.

are 3 DWs at least twice this average, one at the lip and two at the shoulder, and these are considered large enough to signal segmentation points.

According to area proportion, only the large DW (abbreviated LDW) at the lip could yield a sufficiently small subpart, in this case a neck assembly. The shoulder LDWs yield area proportions within the normal body limits, namely greater than 30% as seen from both top and bottom. This is misleading, however, because the shoulder LDWs are below the actual body-neck junction. Their area proportions from the top are thus swelled by including part of the body.

The program is aware of this possibility for both foot and neck LDWs. It seeks out the junction point above the shoulder LDWs (this process is explained below), and it finds that indeed the area proportion is less than 30% from this point. Thus the shoulder LDWs are chosen over the lip LDW for guiding segmentation of the neck assembly.

PINNING DOWN THE PRECISE SEGMENTATION POINT

The inclination of the contour portion within the higher of the shoulder LDWs is obtained by drawing a straight line between the end points of the interval and by calculating the angle ALPHA this line forms with the x-axis. The consecutive intervals above the higher shoulder LDW also have their inclinations computed until one is found that is steep enough to indicated the precise segmentation point. A steep inclination is one that satisfies either of the following criteria:

1. inclination > 85°

2. inclination > 45° and
   inclination > 2 * ALPHA

In the figure, the interval immediately above the shoulder LDW (this interval has a DW of 2) satisfies the second criterion. Hence the precise segmentation point is the lower end point of this interval.

The rationale for the first criterion is that an ideal starting point for a new cylinder is an inclination of 90°, which would give the cylinder vertical sides. The requirement is reduced from 90 to 85° for error tolerance. The rationale for the second criterion is that an inclination twice that of the LDW interval is a significant enough difference to be noted. A 45° lower limit is imposed because a nearly horizontal LDW interval would still yield a nearly horizontal inclination upon doubling its ALPHA. The 45° limit represents a compromise between going straight up, yielding zero width change and a perfect cylinder, and going straight across, yielding infinite width change and an ideal segmentation point.

The search for the segmentation point is conducted differently according to whether the contour portion in the LDW interval slants in or out as seen from the bottom. When the contour slants in, the search for a segmentation point occurs above the highest point of the interval. When slanting out, the search occurs below the lowest point (the location of angles in the second quadrant requires a slight change in computation). In figure 3.1, the contour within the upper LDW interval slants in; hence the outline is examined above to arrive at the indicated neck segmentation

point.

SETTING THE INTERVALS ALONG THE AXIS

A useful interval size divides the axis into roughly 25 steps. This represents a compromise between too many steps, making the program subject to small variations of contour, and too few, blurring out essential features. These adverse effects, nevertheless, may be present with any choice of interval size, and suitable measures must be devised to detect their occurrence.

Small variations in contour may yield LDWs by fortuitous placement of intervals. This situation may be detected by using a larger interval size and by matching the resulting LDWs against those generated from the smaller interval size. If a contour portion yields an LDW under both interval sizes, its LDW is presumed significant; otherwise, the LDW is discarded. The program uses an interval 3/2 the smaller to carry out this check. All three LDWs of figure 3.1 survive this test.

Fortuitous placement of intervals may also mask contour portions that would have yielded LDWs with a slightly different placement. One way of detecting this situation is to interleave another set of intervals with the first placement, such as by coinciding the boundaries of one interval set with the midpoints of the other set (Berthold K. P. Horn pointed this out to me). Unfortunately I did not do this. Instead, I relied on the DW computations with the larger interval size to point out missed LDWs. These LDWs, when found, were also subject to a significance check by using an

even larger interval size.

This process of adjusting interval size as needed is reminiscent of the Warnock algorithm [Warnock 1969], designed originally for hidden line removal but potentially useful as a general technique for picture processing. The direction of focus here however goes in the direction of smaller to larger intervals, whereas Warnock's algorithm subdivided larger squares into smaller ones.

DETECTING SMALL LIPS AND FEET

Small lips or feet that do not yield LDWs are detected by examining DDWs. Analogous to the DW examination, the large DDWs (abbreviated LDDWs) are those twice the average. Problems with interval placement are if anything worse with LDDWs than with LDWs. LDDWs are easily missed by unfortunate interval placement, and are sensitive to interval size as well. Because of the latter reason, LDDWs that are found as before with two interval sizes are unioned instead of intersected.

Returning to the vase in figure 3.1, a foot was not found while examining DWs. There are four LDDWs of value 2, three at the shoulder and one near the base, that might signal a foot. With the 20% area limitation, only the LDDW at the base qualifies as the foot-body junction. Because the LDDW occurs at a concave contour portion, the precise segmentation point lies at the common boundary of the two intervals yielding the LDDW. If the contour portion were convex, the segmentation point would have been the top of the higher interval.

THRESHOLDS

Like most vision programs, this segmentation program contains various thresholds to tune its performance. An example of such a threshold is the compromise choice of 25 steps per outline. The segmentation program has 9 thresholds.

3.2 Prototype Selection

The program assigns one of the 8 prototypes cylinder, cone, ovoid, bowl, bicone, bell, calyx, and pear. The first four are much more common in the pottery domain than the other four.

THE CONTOUR IS BROKEN INTO CONCAVE-CONVEX SEGMENTS

The 8 prototypes are grouped into 3 classes that reflect the number of convex-concave segments from their contours.

<u>convexity</u>	<u>prototypes</u>
convex or concave	cylinder, cone, ovoid, bowl, bicone
convex-concave	bell, calyx
convex-concave-convex	pear

An unknown shape is assigned to one of these classes according to its contour convexity. The final prototype assignment is made within each class on the basis of mouth and base width, height-width ratio, and other contour descriptors.

When a contour is broken into concave and convex segments, the convex segments are maximized. Relatively straight portions of the contour that border a convex segment at one end and a concave segment at the other are added to the convex segment. Convex segments tend to indicate a body, while concave segments indicate junction or transition. Thus it is desirable to maximize the body extent and minimize the junction extent.

Each segment is then examined for significance: if very low in height compared to the shape height (see HEIGHT in Table 3.1 at the end of this section), the segment is ignored. Body-foot and body-neck junctions often

yield such segments, which must be ignored because they are junction artifacts. Very low segments in the middle of a contour are ignored because they represent minor detail.

Conceivably some shapes might survive the significance test with more than 3 segments, a situation for which there is no class. With a step size of about 12 points of body contour (about half the total height), however, this situation is unlikely to arise.

The descriptive terms in the following prototype delimitations are defined in table 3.1, except for the contour descriptors straight, carinated, and curved, which are left for the next section.

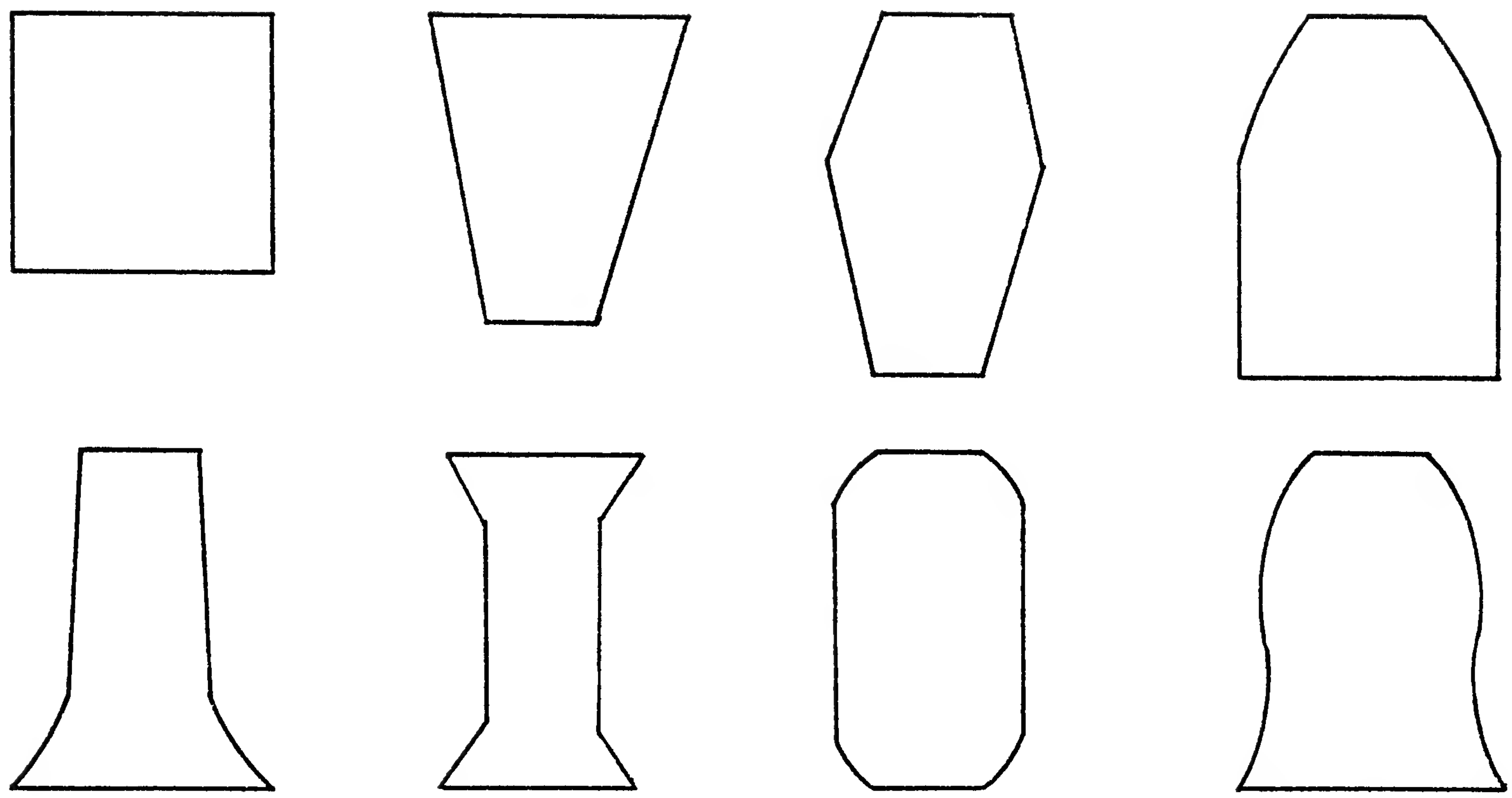
1. *CYLINDER*. A body is a *cylinder* if either
   (1) a high contour portion is straight and vertical (figure 3.2A);
   or
   (2) the contour is concave and vertical (figure 3.2B).

2. *BOWL*. A body is a *bowl* (figure 3.3A) if
   the body is short,
   the mouth is very broad, and
   the body does not satisfy the cylinder or inverted cone
   (figure 3.3B) definitions.

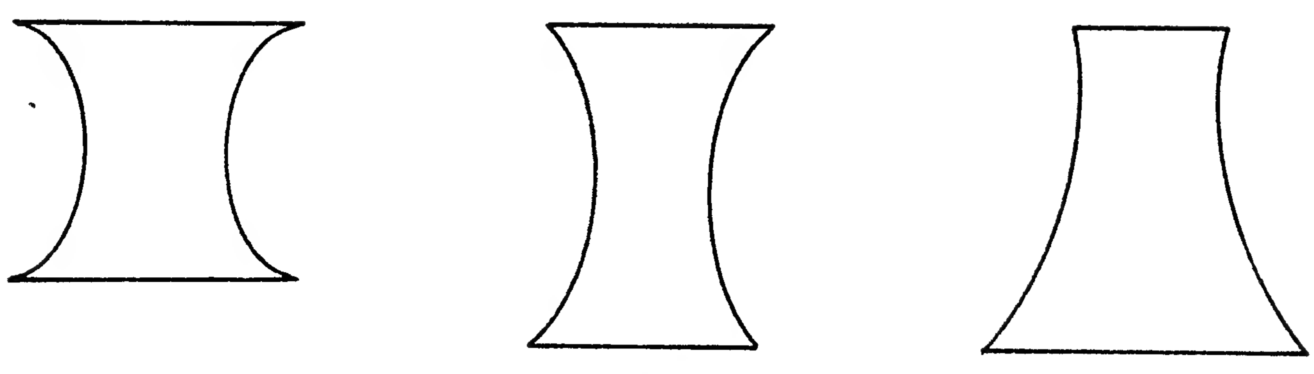
3. *CONE*. A body is a *standard cone* (figure 3.4A) if either
   (1) the contour is straight or concave,
   the contour slants in and is not vertical, and
   the mouth is narrow or broad but not very broad; or

   (2) the contour is convex curved or carinated,
   the contour has a high and straight top portion, and
   this top portion slants in (figure 3.5B).

   A body is an *inverted cone* if
   the contour is straight or concave,
   the contour slants out at a high angle, and



A.



B.

FIGURE 3.2. A. Cylinders with a major contour portion vertical.
B. Cylinders with a vertical, concave contour.

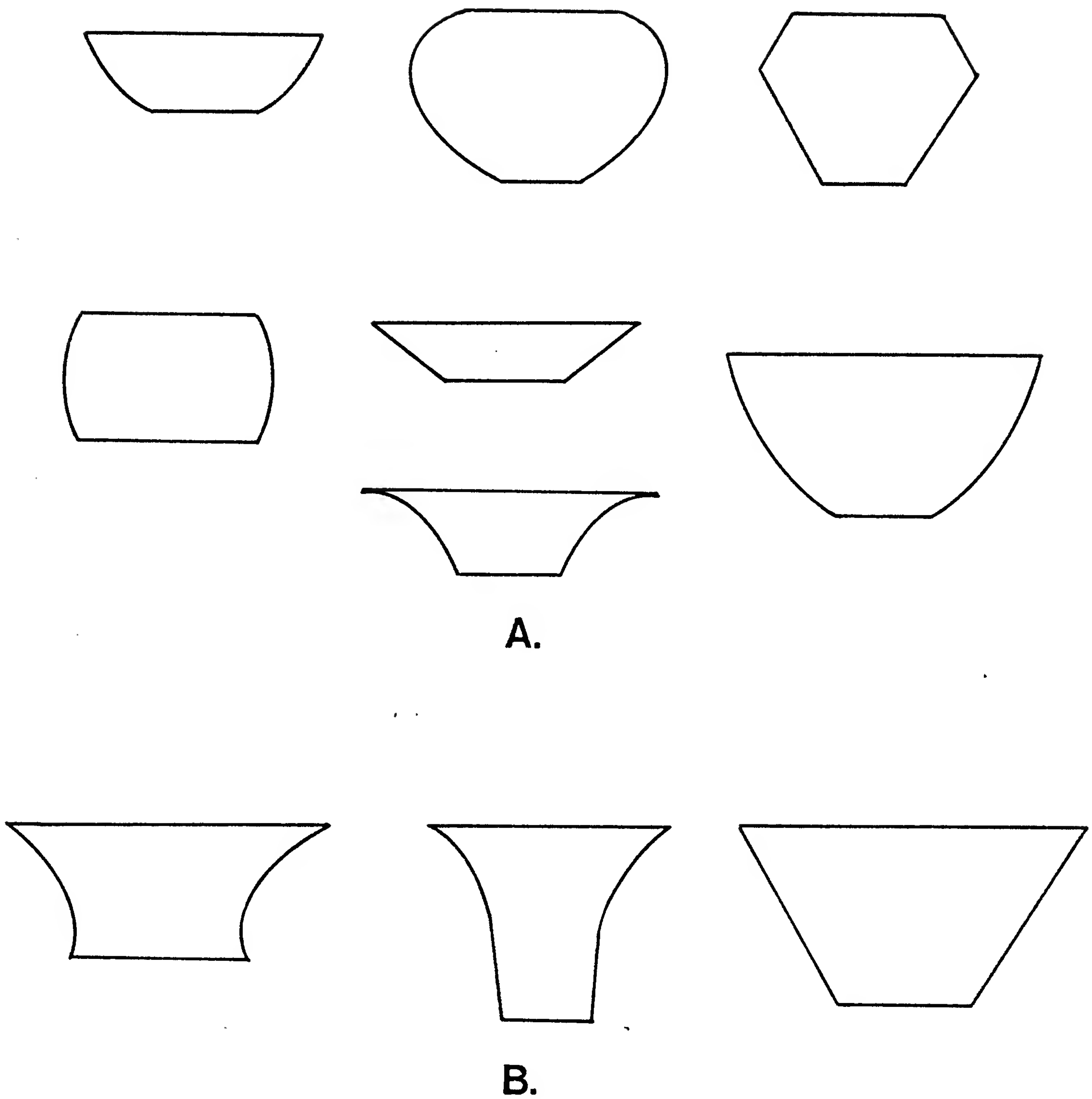
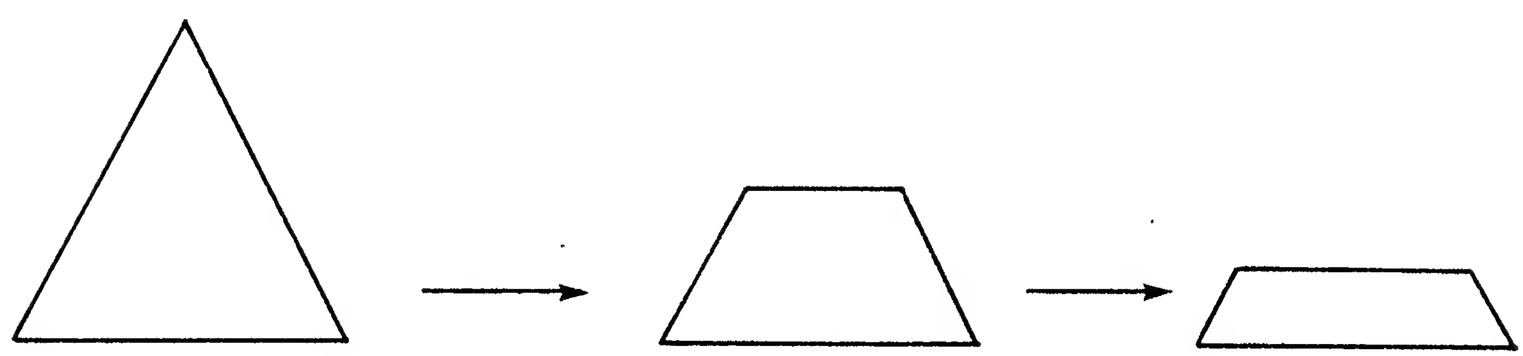


FIGURE 3.3. A. Examples of bowls.
B. Bowls actually considered to be cones because of contour and slant.

A.



B.

FIGURE 3.4. A. Standard cones with straight and concave sides.
B. When too much is truncated from the point of a cone, it loses its identity.

the base is broad but not very broad.

The second set of criteria for standard cones serves to divide the class of ovoids from the class of cones. The mouth width limitation for standard cones and the base width limitation for inverted cones prevents overly truncated cones from being represented as such. Excess truncation causes a cone to lose its identity (figure 3.4B), and manifests itself through very broad mouths.

Orientation is important in distinguishing cones from bowls. When the top is broader than the bottom, the body looks like a bowl. When the bottom is broader, it looks like a cone. Thus if a bowl is turned upside down, it becomes a cone. Inverted cones are exceptions to this rule.

A possible explanation for this rule is found in Arnheim's observation that people view objects by looking from the bottom up. The important feature of a cone is that its sides converge to a point. When the top is narrower than the bottom, the sides tend to converge to a point while scanning upwards. This yields a cone interpretation. When the top is wider than the bottom, the sides appear to diverge. The top appears open, which is the distinguishing feature of bowls.

Though the sides diverge, they may be close enough at the base to appear to have originated from a point. This gives rise to inverted cones. The requirements on contour and base width is an attempt to define when it is that the bottom appears pointlike. Very straight or concave sides facilitate the ability to see the bottom as pointlike, while convex curved sides make the bottom appear rounded as for a sphere truncated at the

bottom. The slope is important because a low angle makes the body look too flat to be a cone.

4. ***OVOID.*** A body is an *ovoid* if
   the contour is convex curved (figure 3.5A), and
   the body is not a bowl, cone (figure 3.5B), or cylinder (figure 3.5C).

5. ***BICONE.*** A body is a *bicone* (figure 3.6A) if
   the contour is carinated, and
   the body is not a bowl, cylinder (figure 3.6B), or cone (figure 3.6C).

This prototype receives its name from its carinated sides, which give the appearance of a standard cone placed on an inverted cone.

6. ***BELL*** and ***CALYX.*** A body is a *bell* or *calyx* if
   the contour is concave-convex,
   the narrow portion is at the convex end while the wide portion is at the concave end (figure 3.7A and B),
   the height-width ratio is approximately one (figure 3.7C), and
   the junction point of convex with concave does not form a local minimum in width (figure 3.7D).

These two shapes are closely related, and are distinguished only by the extent of the convex portion of the contour relative to the concave portion. If the concave portion is the major portion, the body is a ***calyx***; otherwise it is a *bell*.

7. ***PEAR.*** A body is a *pear* if
   the contour is convex-concave-convex,
   one end is very narrow and the other is very broad (figure 3.8A), and
   the body does not have a minimal width point (figure 3.8B).

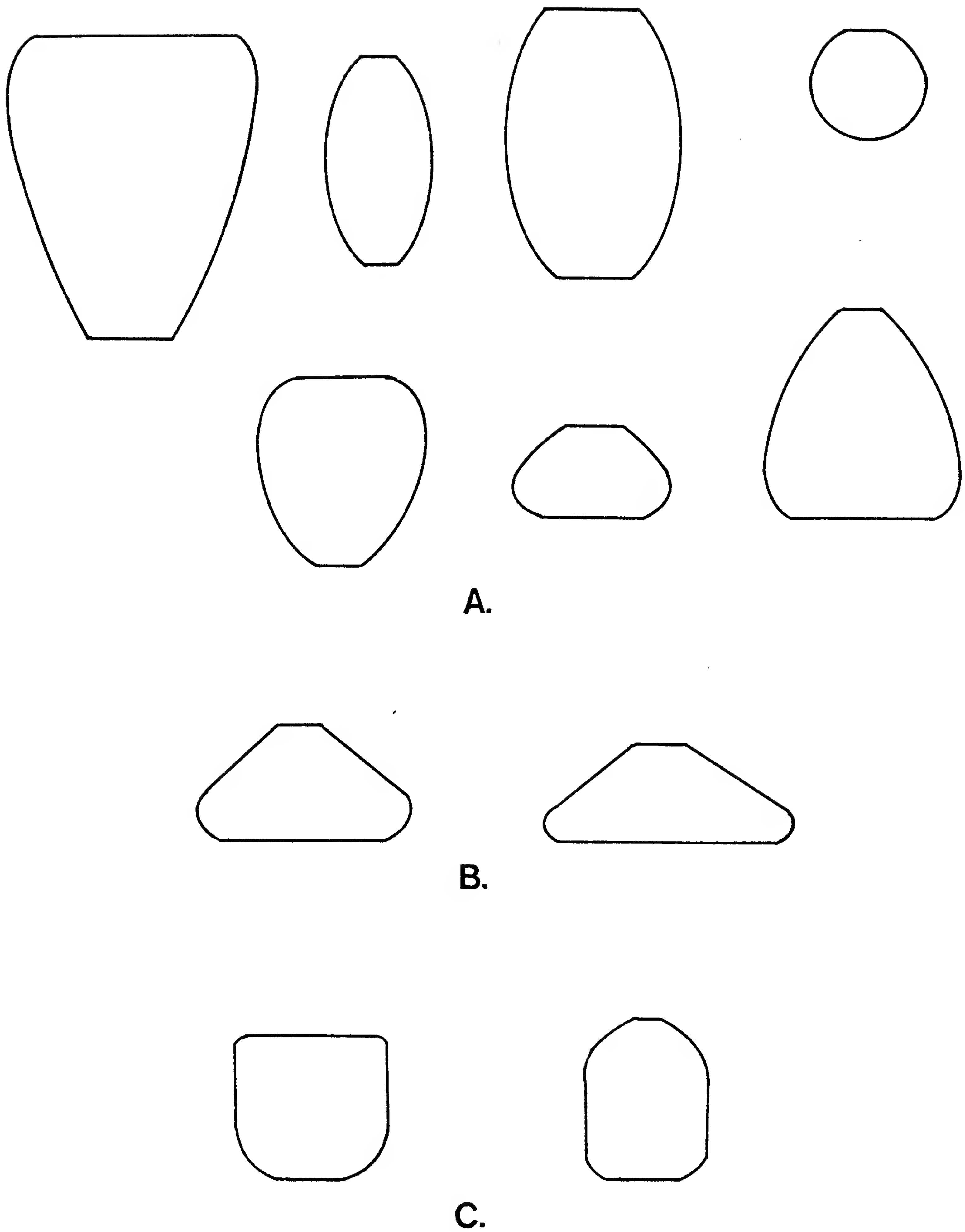


FIGURE 3.5. A. Examples of ovoids.
B. Ovoids are interpreted as cones when the top portion is straight and slanted in.
C. Ovoids are interpreted as cylinders when a major portion is straight.

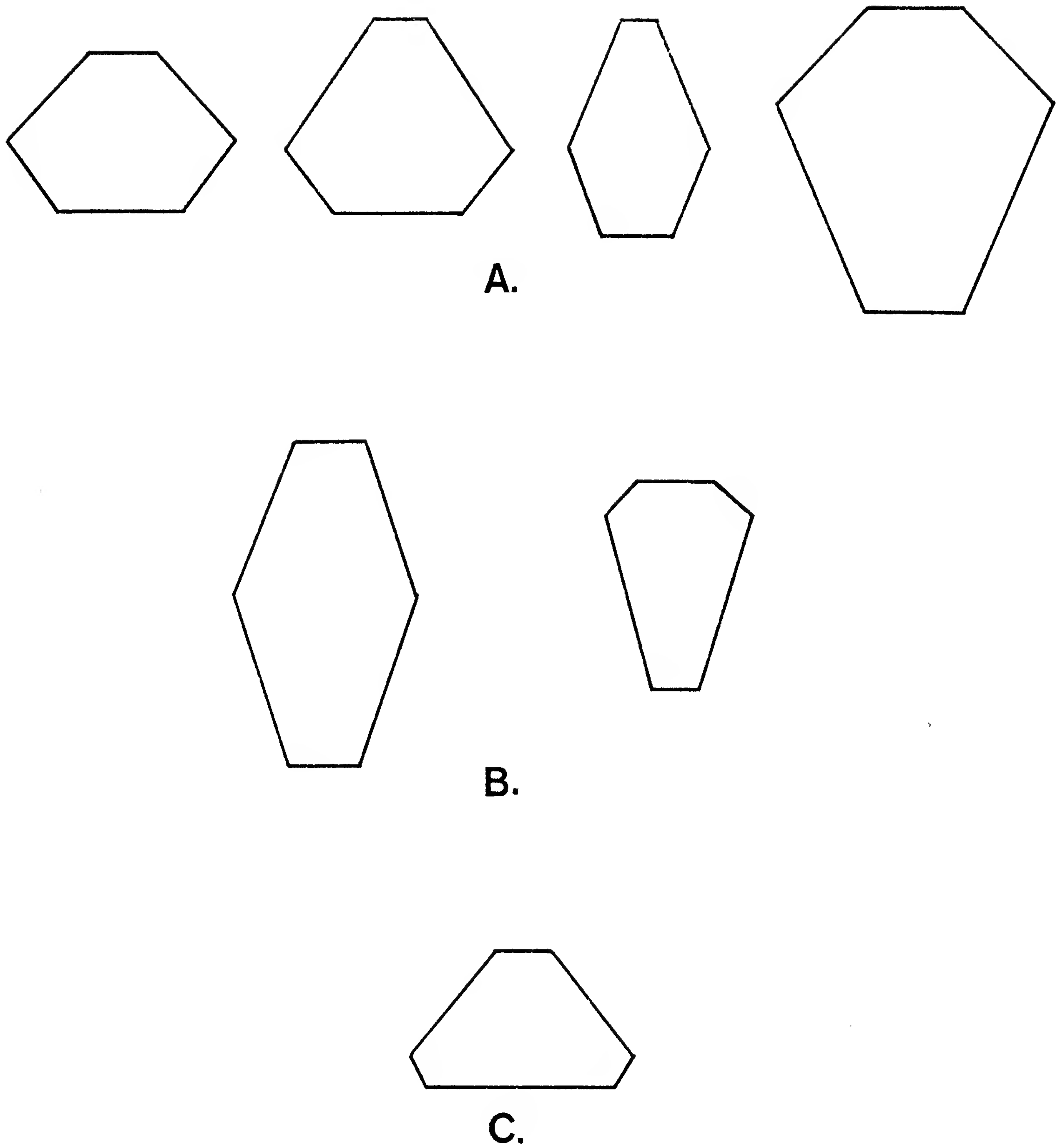


FIGURE 3.6. A. Examples of bicones.
B. A bicone is interpreted as a cylinder when a major portion is straight and vertical.
C. A bicone is interpreted as a cone when the carination part is very low and the sides slant in.

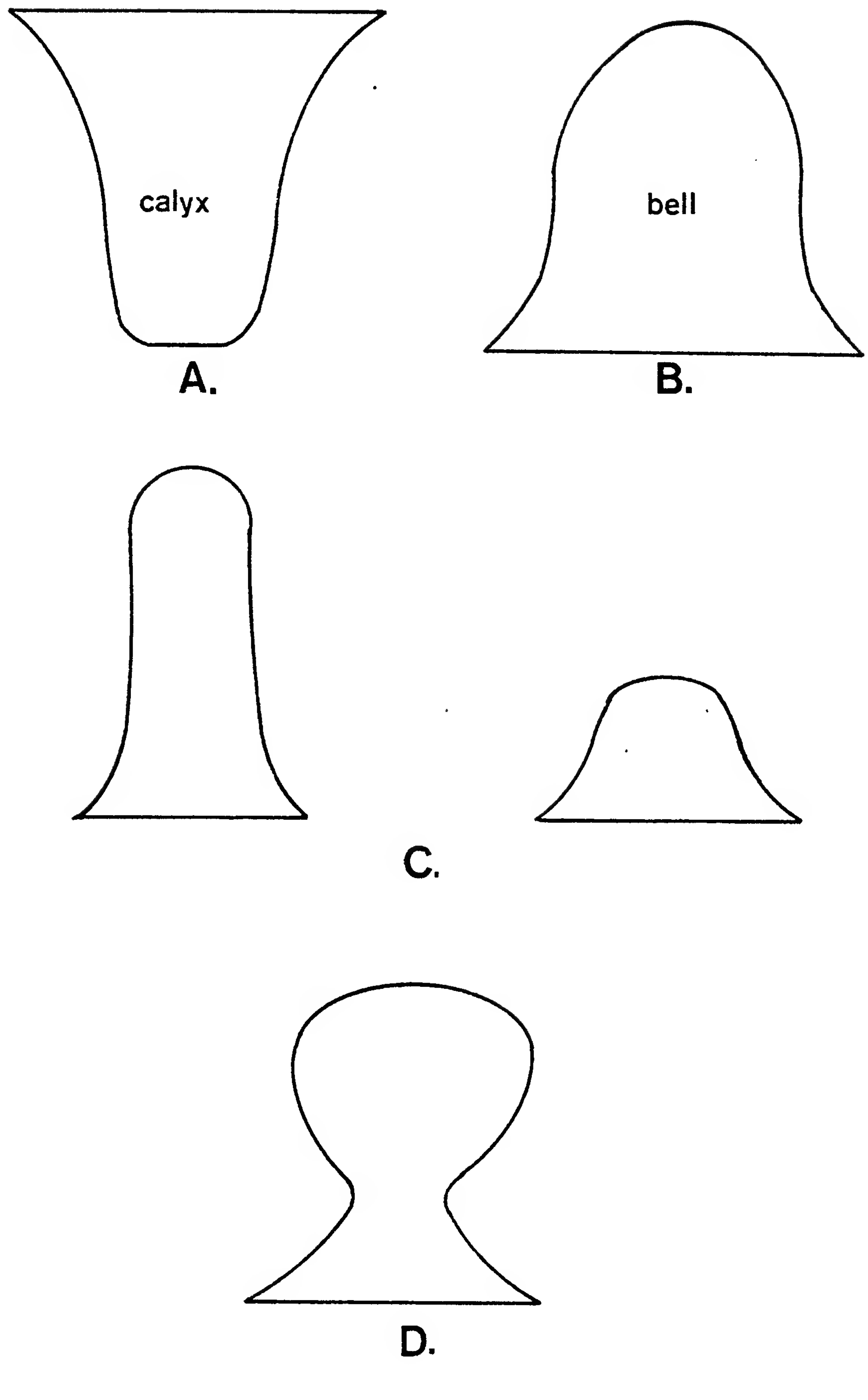


FIGURE 3.7. A & B. Prototypical calyx and bell shapes.
C. Height and width must be about the same.
D. The shape must not have a minimal point.

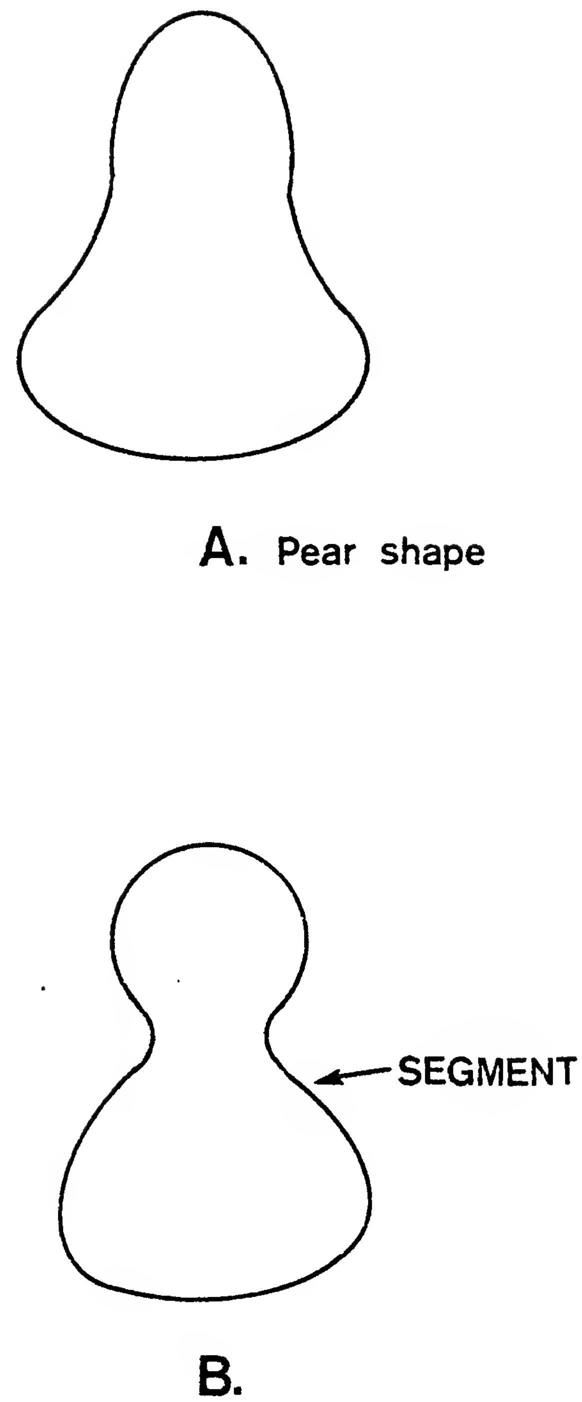


FIGURE 3.8. A. A prototypical pear shape.
B. A pear shape cannot have a minimal point.

SOME BODIES ARE RESEGMENTED IF PROTOTYPE MATCHING FAILS

Prototypes will be successfully assigned to all bodies with contours of one convexity, but the available bell, calyx, and pear prototypes will not cover all bodies with more complex contours. In the latter circumstance, the body is simplified by further segmentation. A convex-concave contour is segmented at the convex-concave junction point. A convex-concave-convex contour is segmented at the junction point of the largest convex segment with the concave segment (figure 3.8B). The concave-convex-concave case does not normally escape the initial vase segmentation.

The two resulting parts are interpreted in a domain dependent manner. Often the bottom part is added to the foot to yield a large pedestal or stand. Less likely, the top portion is added to the neck; for, necks are seldom ornate and do not attain the size of pedestals or stands. A final possibility, not incorporated into the present program, is to describe the body in terms of two prototypes.

PROTOTYPES AND AESTHETICS

It can be argued from aesthetics or simplicity criteria that the three major prototypes cone, cylinder, and ovoid are universal. If one assumes convexity, straightness, gradual curvature change, and slope or verticality are the essential primitive shape descriptors, these prototypes are the simplest in terms of them. Birkhoff [Birkhoff 1933] has also argued that these parameters or their equivalents serve as the basis for aesthetic

judgments of vases. The curvature of an ovoid changes gradually from straight at one end to strongly curved at the other. It should be noted that an ellipsoid has the least curvature in the middle and the greatest at the ends, making it perhaps more complex and less desirable a prototype than ovoid.

Table 3.1

MOUTH: The horizontal straight portion at the top.

BASE: The horizontal straight portion at the bottom.

WIDTH: Let x be the ratio of the mouth or base width to the maximum body width. Then the width descriptors are:

x < 0.1	extremely narrow
x < 0.2	very narrow
x < 0.4	narrow
x > 0.4	broad
x > 0.6	very broad
x > 0.8	extremely broad
x > 0.95	open

HEIGHT-WIDTH: Let x be the ratio of the body height to width. Then the height-width descriptors are:

x < 0.25	extremely short
x < 0.5	very short
x < 1.0	short
x > 1.0	tall
x > 1.5	very tall
x > 2.0	extremely tall

HEIGHT: Let x be the ratio of the vertical extent of a contour portion to the body height. Then the length descriptors are:

x < 0.125	extremely low
x < 0.25	very low
x < 0.5	low
x > 0.5	high
x > 0.75	very high
x > 1.0	extremely high

SLOPE: Let THETA be the angle a straight line from the beginning to end of a contour portion makes with vertical. Then the descriptors for THETA are:

THETA < 15 degrees	vertical
THETA < 45 degrees	high angle
THETA > 45 degrees	low angle

3.3 Modifier Assignment

Once a prototype is selected, it is modified to conform more exactly to the body shape. Modifications like height-width ratio and contour are general to all prototypes, others like orientation are prototype specific. Though certain modifiers are general, their application is different for each prototype. Thus a cone and cylinder with the same height-width ratios will be assigned different descriptors, because the standard of tallness is different for each. The exact terms may also be different for each prototype; shallow bowl, squat ovoid, and short cylinder all have the meaning "short prototype".

1. HEIGHT-WIDTH

The height-width ratio is coarsely broken into two levels: short and tall. Each of these levels may be further refined, such as short into very short and very, very short. The submodifiers *very,very* are normally replaced by the equivalent submodifier *extremely*.

The assignment of terms to height-width ratios are listed for each prototype in table 3.2 on the next page. Note that the direction of refinement tends towards description of the extremes. There are no specific descriptors for the middle range, which is described instead by negating the extremes; an object might be described as *short but not very short*, or *very short but not extremely short*. This is in correspondence with human usage.

Table 3.2

Descriptors assigned to height-width ratios

<u>BOWL</u>:	ratio < 0.25	very shallow
	ratio < 0.5	very shallow
	ratio > 0.5	deep
	ratio > 0.75	very deep
<u>CONE</u>:	ratio < 0.25	very short
	ratio < 0.5	short
	ratio > 0.5	tall
	ratio > 0.9	very tall
<u>CYLINDER</u>:	ratio < 0.5	very short
	ratio < 1.0	short
	ratio > 1.0	tall
	ratio > 1.8	very tall
<u>OVOID</u>:	ratio < 0.6	flat
	ratio < 0.85 & sum < 0.9	squat
	ratio < 1.0 & sum < 1.0	globular
	ratio < 1.3	tall
	ratio < 1.8	slim
	ratio > 1.8	very slim

The ovoid prototype shows the most specialization in terms. The generic term *very tall* is replaced by slim, *extremely tall* by very slim. A short ovoid is *squat*, a very short one is *flat*. *Globular* is intermediate to generic tall and short. It exists because *sphere* is an important special case of ovoid, and because a sphere is neither tall nor short. The term *globular* is preferred to *spherical* because it places less stringent requirements on the contour.

For the ovoid prototype only, truncation of top and bottom must be considered in the assignment of tallness. This is coarsely done by summing the width ratio of the mouth to the maximum width of the body with the corresponding base ratio. The larger the sum, the greater the amount of truncation. This sum is also represented in the table.

2. CONTOUR

Contours of one convexity are assigned one of the three curvature terms straight, curved, or carinated. Each curvature term receives additional refinement:

very straight
fairly straight

gently curved
rounded or circular
strongly curved

slightly carinated
carinated
sharply carinated

Curvature of a line is measured relative to a standard curvature value

that is obtained from a half circle represented by n equidistant points (figure 3.9A). The standard curvature is the difference in angle between two neighboring segments, 180/(n-1). A given curved line is broken into n points, and the curvature between each two segments is calculated. The average curvature is compared against the standard, and is quantified as follows:

average	maximum	descriptor
< 0.25 standard	11 degrees	very straight
< 0.5 standard	11 degrees	fairly straight
< 0.75 standard	20 degrees	gently curved
< 1.5 standard	30 degrees	rounded
> 1.5 standard		strongly curved

A maximum is placed on any one curvature between segments to insure that the line does not curve too much at one point, even though the average is within limits.

A complicating factor is truncation caused by a neck, lip, or foot. It would be incorrect to calculate the standard curvature as 180/(n-1), since n-1 points are being placed on something less than a half circle. Compare for example the two vases in figure 3.10. Both have similarly curved contours, but one vase has a much wider mouth and base than the other. Yet the contour of the squat vase is actually much more strongly curved than the contour of the tall vase. If the truncation width is a and the maximum width is r, then the half circle is reduced by arcsin (a/r) (figure 3.9B). The final reduced circle is divided by n-1 to give the standard curvature.

When a significant portion of a curved line (at least 1/4th the

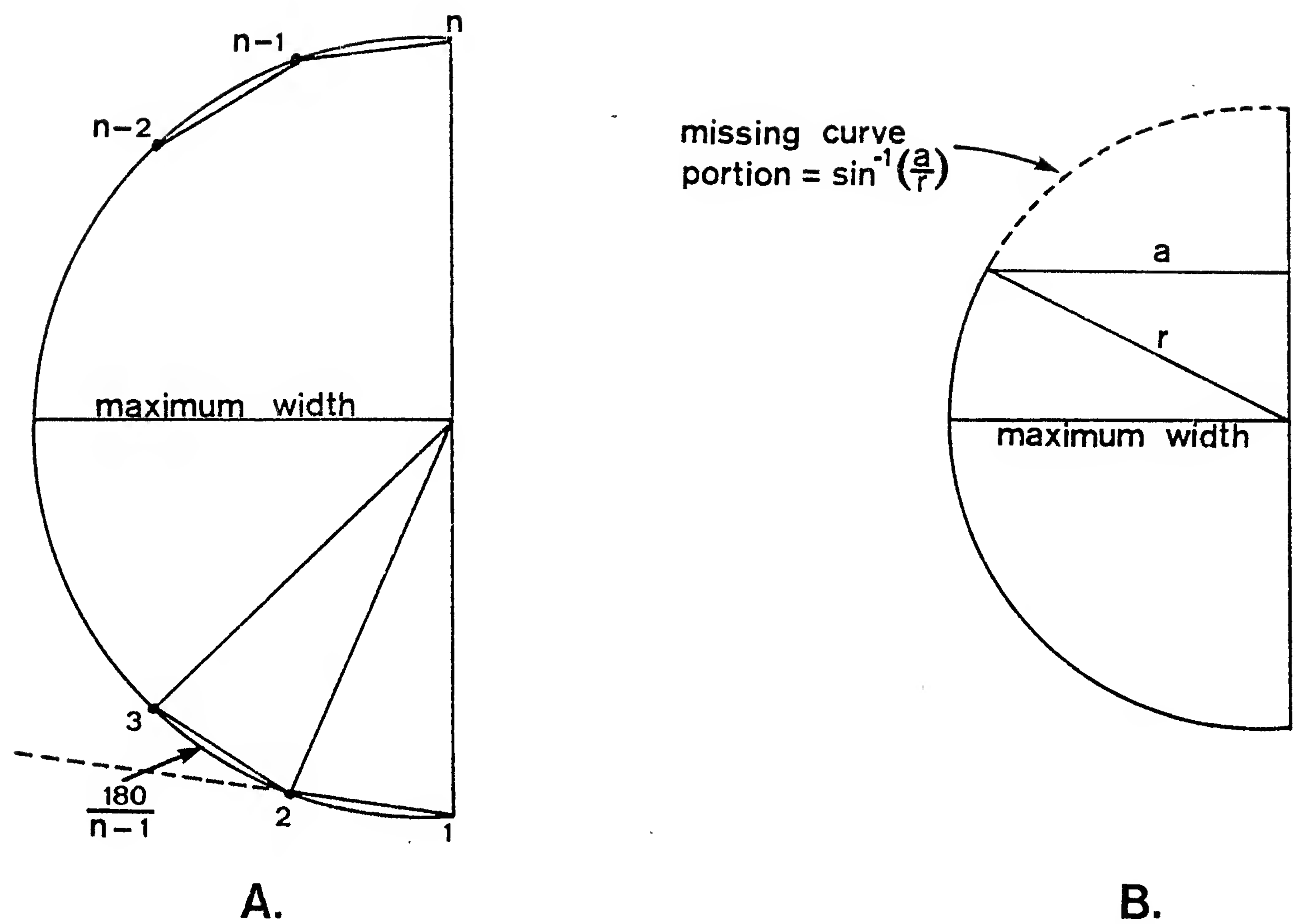

n
n−1
n−2
maximum width
3
$\frac{180}{n-1}$
2
1
A.
missing curve
portion = $\sin^{-1}(\frac{a}{r})$
a
r
maximum width
B.

FIGURE 3.9.

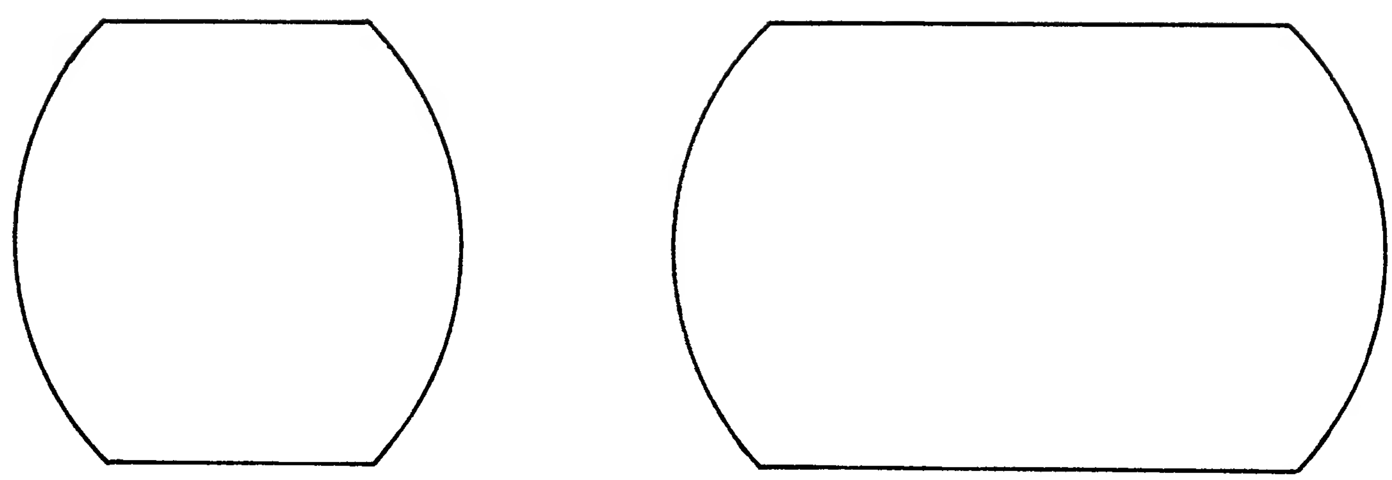


FIGURE 3.10.

length) is straight and lies near one of the ends, a composite curve description is produced as discussed earlier. The straight portion is connected via a quantified transition to the remaining curved portion. Thus a line might be described as *straight becoming gradually rounded.* The transition descriptors are:

direction	descriptor	points
higher	becoming	3 or more
	becoming abruptly	2
	becoming very abruptly	1
lower	becoming	1 or 2
	becoming gradually	3 or more

This table is predicated on a contour of roughly 12 points, and is applied as follows. After the straight portion is split off, the remaining portion is assigned an average curvature. The first point on the curved portion at which this average is reached is located. For example, suppose the average is 12 degrees and the curved portion has the curvature list {10,11,13,14}. If the junction with the straight line is at the 10 degree end, then the average is first reached at 13. The transition portion is thus {10,11}. If the junction with the straight line is instead at the 14 degree end, then {14,13} is the transitional part.

A transition descriptor is assigned depending both on the direction in which the average is approached and on the number of points in the transition. If the average is approached from lower curvature, as when {10,11} is the transition, then the descriptors are *becoming* and *becoming gradually*. If approached from higher curvature, as when {14,13} is the transition, the descriptors are *becoming*, *becoming abruptly*, and *becoming*

very abruptly. The curvature of the curved portion is recomputed after removing the transitional portion.

Finally, the straight portion of the composite curve is assigned a height position on the body: it is on either the lower or the upper profile. The description of the contour of the amphora in figure 3.1, for example, is ***straight lower profile becoming abruptly rounded.***

The degree of carination is computed from the angle between the two straight segments. Its quantification is:

angle < 120 degrees	sharply carinated
angle < 140 degrees	carinated
angle > 140 degrees	slightly carinated

3. CONVEXITY

The convexity of a contour is either convex or concave. If concave, additional descriptors are computed in conjunction with the average curvature of the contour.

contour	convexity
gently curved	slightly flaring
rounded	flaring
strongly curved	widely flaring

4. SHOULDER

All prototypes except cone may have a shoulder. A shoulder exists if:

the top contour portion slopes in,
the mouth is not extremely broad, and
this contour portion is low in height.

A shoulder slopes in to constrict a body's opening. Requiring a not extremely broad mouth has the effect of insuring the opening is constricted

enough for the shoulder to be noticeable. A shoulder's height is low because the shoulder's dimensions must be small relative to the body.

5. GREATEST WIDTH

The point of greatest width is specified for bodies with shoulders and for bodies with carinated profiles.

Let $r = \frac{\text{height from base to point of greatest width}}{\text{height of body}}$.

The r values quantize the point of greatest width as follows:

$r > 0.6$	high shoulder
$r > 0.4$	not described
$r < 0.4$	low belly

6. SLANT

Slant is assigned only to cylinders. If a straight line drawn from mouth to base is within 5 degrees of vertical, the slant is vertical. If the mouth is wider than the base, the contour slants out; otherwise it slants in.

7. ORIENTATION

Orientation is assigned only to cones. When the top is broader than the bottom, the orientation is inverted, otherwise standard.

3.4 Foot, Neck, Lip

The foot and neck assemblies may require further segmentation. A pedestal foot is separated into stem and base. A neck and lip are sought from the neck assembly, though only one of them may be present.

1. FOOT

A foot is assigned one of the prototypes cone, cylinder, pedestal, or molded. The prototypes pedestal and molded are specific to foot. If a foot is absent, the flat base receives only a width descriptor as in table 3.1.

A pedestal is segmented by looking for large DWs, as in section 3.1. Since the stem is much narrower than the base, the relevant large DWs are the ones sloping in. Stem size restrictions allow an intelligent choice to be made among several large DWs. These restrictions are:

the stem width is at most half the base width,
the stem width is narrow relative to the body, and
the stem height is at least half the base height.

The widths of the stem and of the base are described individually. The base width revives the descriptors in table 3.1 The stem width r relative to the body is expected to be rather narrow, and is therefore quantized differently:

$r < 0.125$	very narrow
$r < 0.25$	narrow
$r > 0.25$	broad

The pedestal height relative to the body height is also described, and receives the height descriptors in table 3.1.

Finally, the articulation of the stem-base junction is specified. The junction is ***articulated*** if the contour at that point is angular; otherwise the junction is *splaying*. The degree of splay is computed according to the curvature of the junction contour:

gently curved	slightly splaying
rounded	splaying
strongly curved	widely splaying

The pedestal in figure 3.11 would be described as splaying, high in height, broad stemmed, and narrow based.

If extremely short, the foot is said to be molded, such as the foot in figure 3.1. Because of this extreme shortness, a molded foot does not have a manifested contour. Thus *molded* is not really a prototypical term, but a default category for feet that are too short to be assigned the usual prototypes. The only modifier assigned to molded feet is width. A ***ring*** foot is a particular kind of molded foot with a convex, rounded contour. It is made from a long circular rod of clay wrapped in a circle.

If the foot is not molded, a cone or cylinder prototype is assigned to it as in section 3.2. An example of a cylindrical foot is the psykter in figure 1.1. Contour, convexity, and width modifiers are computed as in section 3.3 and in table 3.1. The foot to body height ratio is quantized as follows:

ratio < 0.125	very low
ratio < 0.25	low
ratio > 0.25	high
ratio > 0.5	very high

FIGURE 3.11.

2. NECK

If a neck assembly is to contain a neck, it must meet minimum size restrictions:

> the assembly is not extremely short, and
> its height relative to the body height is not extremely low.

The body must also have shoulders for a neck to rest on (standard cones, though they may not have shoulders, can sport necks). Presuming the assembly meets these requirements, an attempt is made to segment a lip from the neck. Segmentation is achieved as in section 3.2, with the neck serving the role of the body: the lowest large DW slanting out and forming a subpart of less than 30% area is sought. If such a DW exists, a lip is present. The putative neck contour is finally examined. If it is not roughly cylindrical, there is no neck and the whole assembly is treated as a lip.

Once the existence of a neck has been established, five modifiers to a cylinder prototype are computed for it: height, width, contour, convexity, and slope. In calculating width, the narrowest portion of the neck is used. As a sample neck description, the neck in figure 3.1 is high and broad, with a straight and vertical contour.

3. LIP

Since lips are normally very short in height, most lips fall into the molded category. Now and then lips do arise that require a standard prototype, such as the cup-shaped lip of a lekythos.

Special terms exist for molded lips with certain contour characteristics. A ***rolled*** lip corresponds to a ring foot, and has a rounded convex contour. A lip is ***everted*** if its contour is convex and slants out. A concave lip that slants out is ***flaring***. A ***wide brim*** is an everted lip whose horizontal extent from narrowest to widest point has a ratio of at least 0.1 with the body width.

Possible modifiers for molded lips are width, height, articulation, convexity, and slope. The narrow point of the lip is referred to as the mouth width, which receives the descriptors in table 3.1. The lip height relative to the body is quantified as follows:

ratio < 0.1	very low
ratio < 0.2	low
ratio > 0.2	high
ratio > 0.3	very high

Articulation is determined as for the body-neck junction, and is either offset or not offset.

As sample lip descriptions, the lip in figure 3.1 is rolled, broad mouthed, low, and offset from the neck. The lip in figure 3.11 is very low and broad mouthed.

3.5 Names and Functions

The descriptors in the previous sections facilitate modeling and matching against models. Because these descriptors are relatively free from particulars of the low level input of individual vases, one can concentrate on a vase's relation to general categories such as amphoras.

The program's taxonomy consists of 42 vase names, listed in table 3.3 at the end of this section. Greek pottery dominates the list, because archeologists delineate, depict, and describe this class of pottery more thoroughly than other classes of pottery. Common vases such as bottles and jars constitute the remainder of the list, but I did not develop the taxonomy for these vases as thoroughly as for the Greek vases.

The number of entries in the taxonomy is limited mainly by the difficulty of drawing up adequate specifications for a new vase type. Integrating the new vase type into the taxonomy structure also presents difficulties. A Winston net [Winston 1970] could probably be devised that would automate this addition process.

Strictly speaking, shape in itself is not enough to name a vase. Size and material of construction might cause a bowl, for example, to be variously described as a vat, tub, basin, or cup. Fortunately these attributes do not influence most other names, names that are adequately assigned by shape alone.

ONE OF FOUR FUNCTIONS IS ASSIGNED

Vases are often created for practical use. Accordingly, the program

attempts to assign one of four functions to a vase: solid storing, liquid storing, liquid pouring, or solid-liquid dispensing. The 42 vase types in table 3.3 have been separated by main function. Of course vases may be made to serve more than one function, and vases near the borderline between two functionally distinct categories can serve either function reasonably well.

The basis for assigning function to a vase is the character of the opening. A vase is meant to hold something, and the opening determines what things are easily put in and taken out. When a neck is present, it is relatively difficult to remove material; hence necked vases are storing vases. If the neck is narrow, the vase serves primarily to store liquids. A narrow neck makes pouring easier, and allows transportation with less chance of spilling than does a broad neck. Broad necks are better for getting solids in and out; thus broad necked vases serve to store solids.

Vases without necks serve as temporary containers. When a vase is widest at the opening, it is useful for pouring liquids; examples are cups, bowls, and ladles. When the opening is more constricted, pouring becomes impractical because neck absence and shoulder proximity would cause the liquid to hit the sides. Such dispensing vases include jars and bowls, which more conveniently transport liquids and temporarily store liquids than do pouring vases. Material is removed from dispensing vases by other means than pouring, such as by ladling, picking by hand, or sipping.

DECORATIVE VASES AND NON-VASES

A well-proportioned vase would be assigned one of the previous 4 functions. A vase whose proportions deviate too far from normal could not adequately serve any of these functions. For example, the neck might be too high, too broad, or too narrow; the body might be too tall or short; or the handles might be too delicate. The middle ranges of modifiers indicate the normal proportions expected of a vase. Thus a broad neck is a neck whose width is somewhat greater than the "normal" neck width; a narrow neck is less in width than "normal".

Some misproportioned vases are made for decoration. The name ***vase*** often indicates such a purpose. In current usage, ***vase*** is a flower container with an extremely high neck to accomodate flower stems. Well-proportioned vases are also used for decoration, but the program prefers to assign a practical function if possible.

The submodifier *extremely* indicates that the associated modifier has exceeded the normal bounds. If a vase has too many such modifiers, or if any one modifier is too extreme, it is doubtful that the object should be called a vase at all. Objects which are clearly not vases, but which someone might have entered to fool the program, would not pass the segmentation stage. Implicit in the segmenter is the normal range of size and shape of the vase parts; the segmenter would simply gag on an object that does not fit this mold.

Table 3.3

liquid-storing	solid-storing	pouring-vase	dispensing-vase
bottle	jar	bowl	pot
flask	neck-amphora	cup	krater
florence flask	continuous-curve amphora	pan	column krater
kjeldahl flask	pelike	plate	bell krater
aryballos	stamnos	cooking pot	calyx krater
ampulla	urn	cooking pan	lebes
lekythos		ladle	psykter
oinochoe		mug	
bell-mouthed oinochoe		kantharos	decorative
olpe		kylix	vase
jug		skyphos	
alabastron		kotoyle	
hydria			
kalpis			
pitcher			

3.5.1 Program Structure

The identification program is structured into hierarchical modules. The lower the flow of control in the hierarchy, the more detailed and specific are the shape requirements. A partial listing of these modules and of their main connections is presented in figure 3.12, where one particular line down through amphoras has been detailed. The solid links are considered the normal transitions from a module. Not shown in the diagram is the crosslinking between modules in different parts of the diagram, which are too numerous to depict in this drawing.

A particular module, such as the amphora module, sets forth conditions for a description to fulfill. When it finds a set of descriptors it can key on, the module will either assign a name or pass control to a submodule. The amphora module might, for example, assign the name ***neck amphora*** to a vase. Or the amphora module might decide the descriptors better match the specifications for one of its submodules pelike or stamnos. These submodules represent special kinds of amphoras for which there are distinct names.

A module unable to make an assignment will either return control to its parent module or it will activate a module in some other part of the hierarchy. The Greek jar module, for example, could activate the jar, jug, or pot module if the descriptors warrant it. This requires the Greek jar module to have some knowledge about likely causes of failure and about courses of action to take when they occur.

This program is a sort of generalization of a decision tree approach.

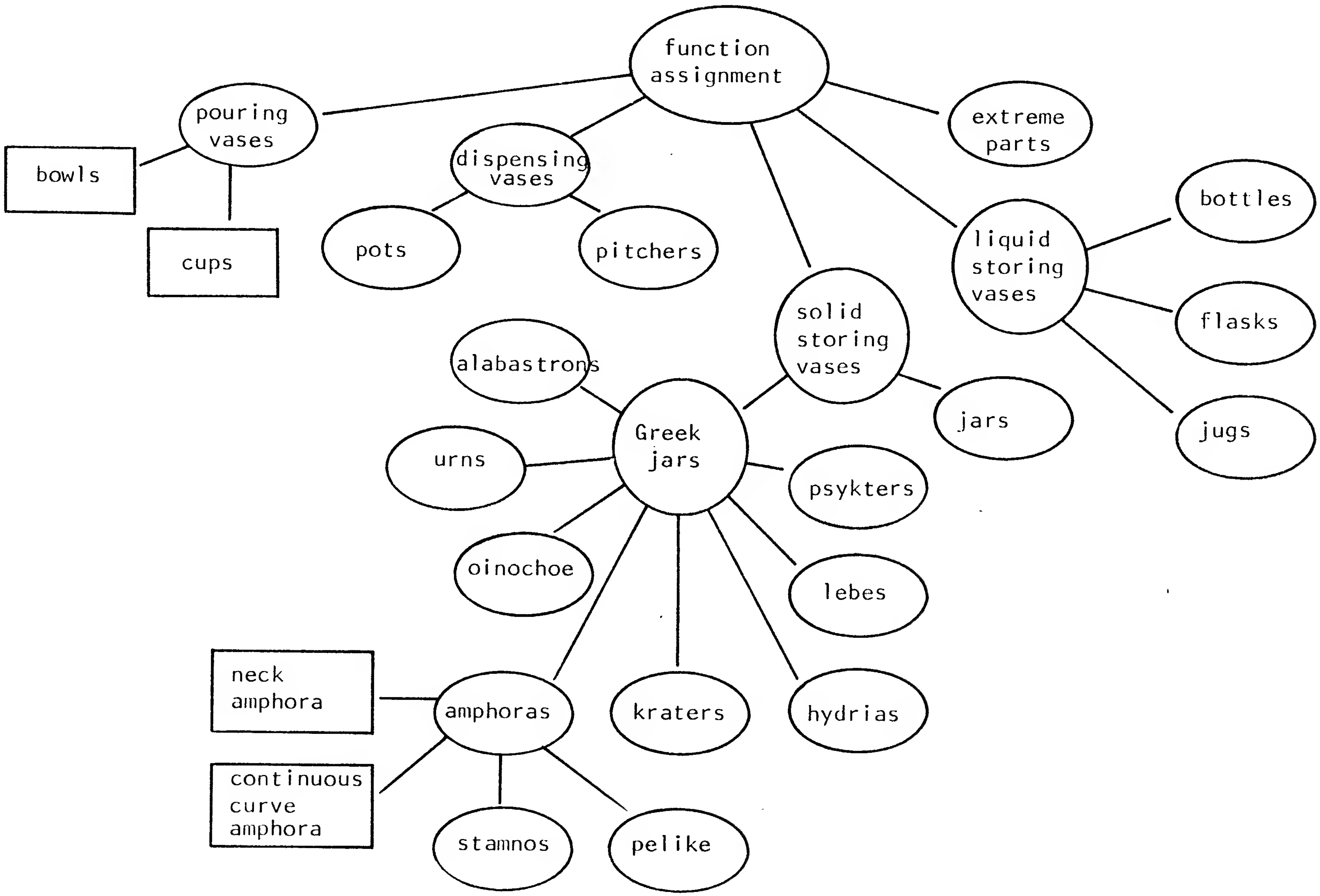

function assignment
extreme parts
pouring vases
bowls
cups
dispensing vases
pots
pitchers
liquid storing vases
bottles
flasks
jugs
solid storing vases
jars
Greek jars
alabastrons
urns
oinochoe
psykters
lebes
hydrias
kraters
amphoras
neck amphora
continuous curve amphora
stamnos
pelike

FIGURE 3.12.

A simple decision tree is a good first rough cut through bewildering variations and exceptions at dividing vases into categories. Categories are groupings of vases based on function and on feature similarity. Vases in a category must cluster closely enough to be distinguishable from vases in other categories. The stronger the clustering, the better is a decision tree approach toward classification, because there are fewer exceptions and less overlap between categories. The hierarchical nature of the shape descriptors itself gives impetus to such a scheme. A distinction between vases with broad or narrow necks, for example, naturally forms two new branches from a node of the tree.

CROSS CONNECTIONS CIRCUMVENT BOUNDARY PROBLEMS

A decision tree is inadequate because cross connections across node levels are needed. The need arises from boundary fuzziness between modifier terms, from boundary fuzziness between prototypes, from fuzziness of part distinctions as between neck and lip, and from diversity of vases in a category. To avoid confusion and to cut down on the number of possibilities, these cross connections are more easily added to a basic decision tree than incorporated at the very beginning of drawing up a taxonomy.

Boundary fuzziness must be considered when vase features lie near a boundary. A tree node might key on broad necks, for example, or on ovoid bodies in order to pass control to a subnode. This subnode may not care that the neck is narrow but nearly broad, or that the body is a cylinder

which has such a curved contour that it is almost an ovoid. Later processing along a different branch of the tree should detect this situation and pass control to this subnode.

Thus one part of the tree must sometimes have some form of model of another part. The decision as to which node has what models is entirely specific to the domain, and depends on what exceptions or variations are likely to reach a node. How far up or down the tree a cross connection is made depends on frequency or importance: the more likely a cross connection situation, the higher in the tree it should be looked for.

Even without boundary problems, a category may have a great deal of latitude in what descriptions satisfy it, especially such general categories as cup, bowl, and jar. The corresponding node in the tree must be accessible from diverse paths.

Finally, with regard to the neck-lip fuzziness, a very low neck may be almost indistinguishable from a high lip. When one is looked for, the other should be expected also.

AN EXAMPLE OF STRAIGHTFORWARD IDENTIFICATION

The amphora in figure 1.2 is a prototypical amphora, and is identified by following the main links of the tree. The program begins with the function module, which assigns to the amphora a solid storing function because of its broad neck. Before this assignment is made, the function module checks if the vase is well proportioned. For misproportioned vases, the module determines if the vase is one of several special types, such as

decorative vases with extremely high necks or ladles with extremely long handles. If no special assignment can be made, the module refuses to recognize the object as a vase and fails.

Control now passes to the solid storing module (figure 3.12), which keys first on body shape. Because the body of the amphora is ovoid, control is passed to the Greek jar module. Necked vases with ovoid bodies are typical of Greek jars, and a higher level module for them is useful.

The prototypical Greek jar has moderately sized foot, neck, body, and lip, and has one or two handles. The Greek jar module treats as special cases those jars that do not fit this prototype, such as a lebes which has neither handles nor neck. Our sample amphora fits the prototype. Handles, a key descriptor, are examined next. Although exact handle shape is seldom important, handle presence or absence goes a long way towards determining vase names. The only difference between an amphora and hydria, for example, is that a hydria has an extra handle. Because the sample amphora has two handles, the amphora module is activated.

The amphora module begins by checking handle orientation. The vertical handles of the sample amphora cause the module to focus on the body, because there is a subtype of amphora called a pelike which has vertical handles and a low belly. The sample amphora has instead a high shoulder. A final distinction is based on body-neck articulation: when articulated, as is the sample amphora, the vase is referred to as a neck amphora; otherwise it is referred to as a continuous-curve amphora.

A note on figure 3.12: names enclosed by ovals designate modules that

do processing. Names enclosed by rectangles are terminal labels that the attaching module may assign.

AN EXAMPLE OF CROSS CONNECTION OCCASIONED BY MODIFIER BOUNDARY

Amphoras and other Greek jars may have narrow necks. The main pathway to the amphora module, however, comes from the solid storing module, which deals only with broad necked vases. In order to identify narrow necked amphoras, a cross connection to this pathway is required.

Let us imagine that the amphora in figure 1.2 has a narrow neck, but is otherwise unchanged. Seeing the narrow neck, the function module passes control to the liquid storing module. The latter module, noting the two vertical handles, activates the jug module, because a typical jug has a narrow neck and one or two vertical handles. The jug module is alerted, however, by the ovoid body with two vertical handles. It knows about narrow necked Greek jars, and makes a cross connection to the Greek jar module. To be sure, a narrow-necked amphora is also a jug, but the program prefers to assign the more specific vase type. The cross connections from the jug module are indicated in figure 3.13 by dashed lines.

A CROSS CONNECTION OCCASIONED BY PROTOTYPE BOUNDARY

The dividing line between the ovoid and bowl prototypes hinges on mouth width. A squat ovoid with broad mouth transforms into a deep, shouldered bowl if the mouth becomes very broad. These two shapes are quite similar, and the krater, for one, finds their separation into

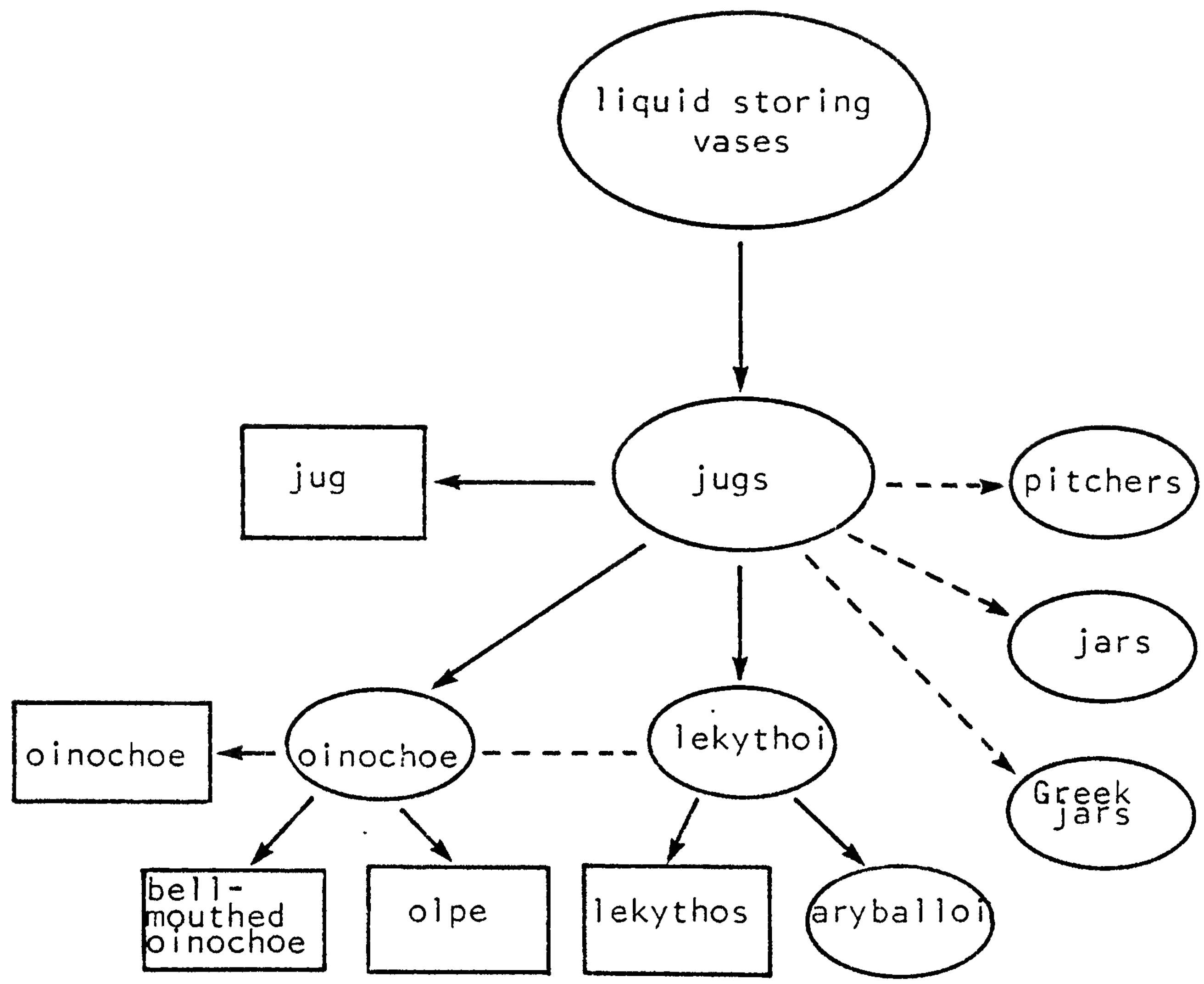

liquid storing vases
jug
jugs
pitchers
jars
oinochoe
oinochoe
lekythoi
Greek jars
bell-mouthed oinochoe
olpe
lekythos
aryballoi

FIGURE 3.13.

different prototypes unimportant.

Kraters cover a broad class of vases. The body may be an exotic bell or calyx shape; or it may be a squat ovoid or a deep, shouldered bowl. A neck may or may not be present; if present, it must be low. The mouth ranges from broad to open. The main pathway to the krater module comes from the Greek jar module. Since the Greek jar module is normally reached by vases with ovoid bodies and broad necks, a krater with a deep bowl body and no neck (figure 3.14) must travel a different path to be identified.

The function module would pinpoint this krater as a pouring vase. The pouring vase module separates vases into two clases: those with handles and those without (figure 3.15). Upon activation, the handled bowl module notes the two vertical handles and the bowl shaped body, which it knows is characteristic of the Greek drinking cup kantharos. A kantharos often has large, high-flung handles (figure 1.1). The kantharos module is alerted, therefore, by the small size of the krater handles. The clincher though is the shoulder with small lip. A kantharos cannot have shoulders unless a wide brim reaches to the limits of the body width; otherwise it is too difficult to drink from it. The kantharos module knows that some kraters are similar in shape to kantharoi save for these characteristics, and activates the krater module.

Figure 3.15 indicates that cup and bowl labels are assigned in diverse modules. There are in fact no separate cup and bowl modules. These two vase types are so varied and pervasive that I was forced to work under the assumption that all pouring vases are either cups or bowls unless proven

FIGURE 3.14.

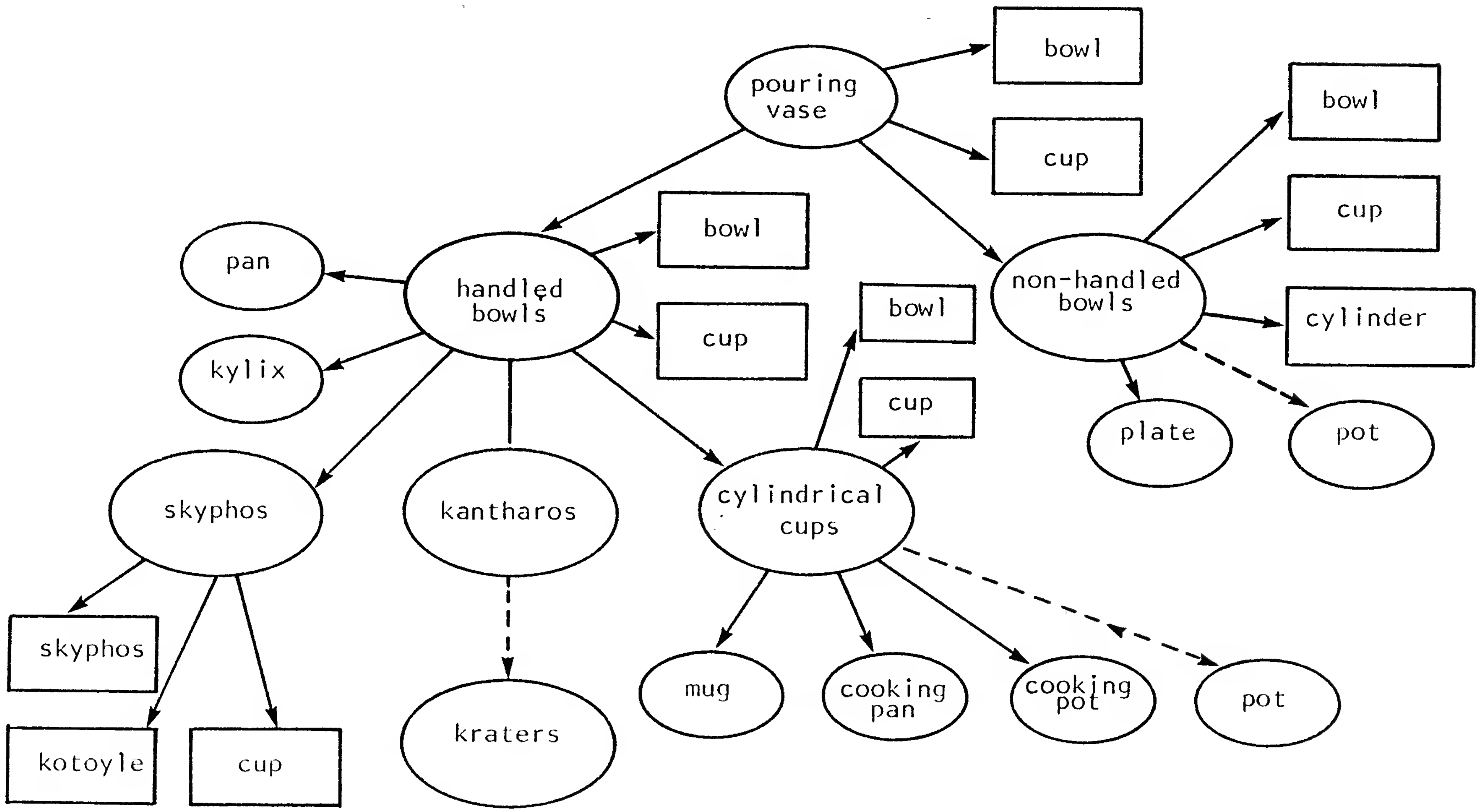

pouring vase
bowl
cup
non-handled bowls
bowl
cup
cylinder
plate
pot
handled bowls
bowl
cup
pan
kylix
skyphos
kantharos
kraters
cylindrical cups
bowl
cup
mug
cooking pan
cooking pot
pot
skyphos
kotoyle
cup

FIGURE 3.15.

otherwise. The cup and bowl labels are in a sense default assignments for a module, while the oval modules are special cases recognized by the parent module.

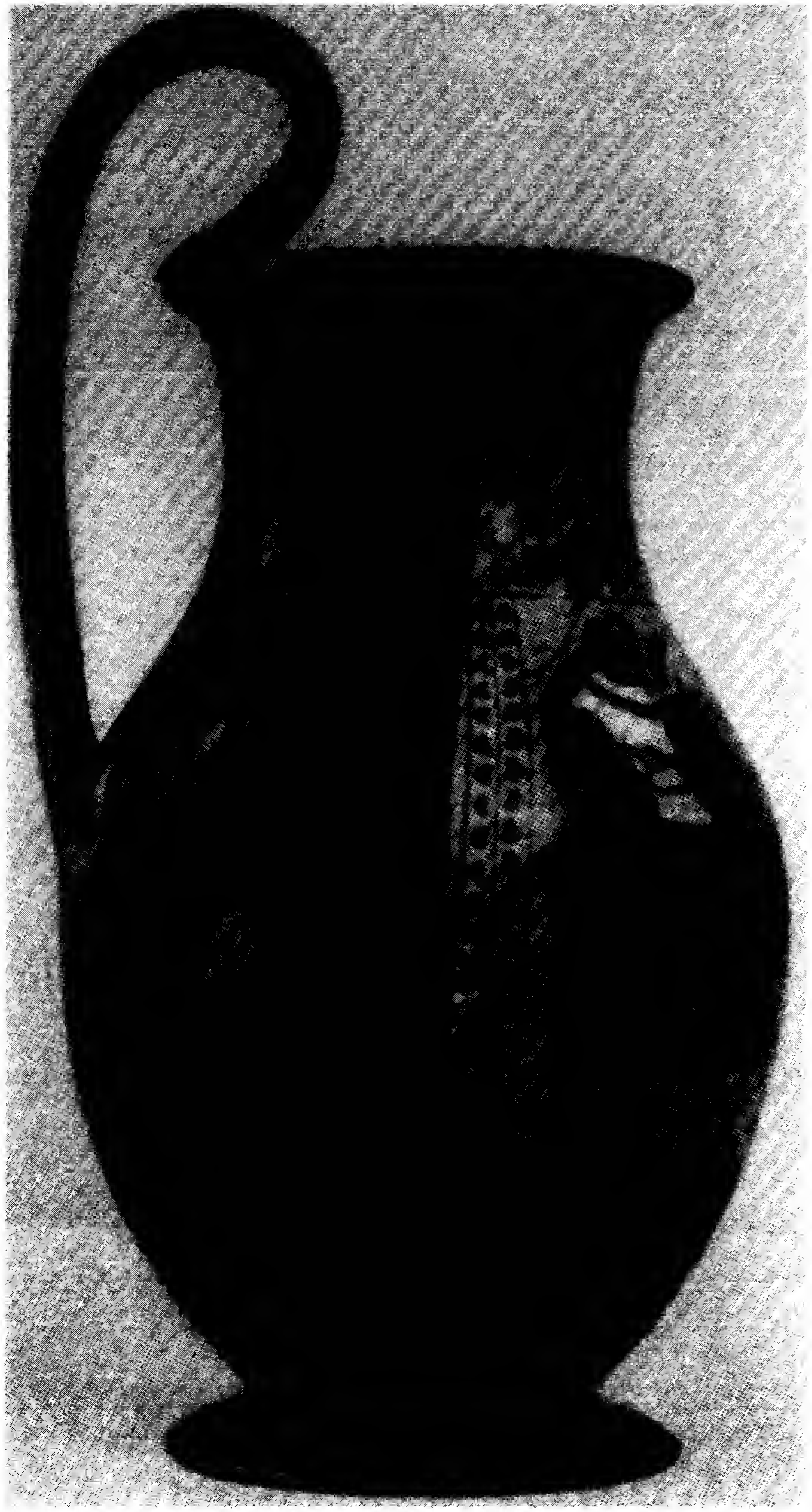
A CROSS CONNECTION OCCASIONED BY PART DISTINCTION

Low necks that are not offset from the body, such as the neck of the continuous curve oinochoe in figure 3.16, may appear indistinguishable from flaring lips. Depending on the interpretation given such necks during segmentation, totally different paths would be followed in the naming tree. If the oinochoe in figure 3.16 were described as having a flaring lip but no neck, a cross connection to the normal oinochoe pathway in figure 3.13 is required.

This oinochoe would be assigned a dispensing function by the function module. The dispensing module has many cross connections to other modules (figure 3.17) because of the fuzzy line between low necks and lips. The dispensing module knows that the jug module is unconcerned about this distinction, and so when the dispensing module sees in the oinochoe the jug characteristics ovoid body, single vertical handle, and narrow mouth, it calls on the jug module. The jug module continues by activating the oinochoe module. The oinochoe module knows about the low, non-offset neck vs. flaring lip confusion, and successfully identifies the oinochoe.

A CROSS CONNECTION OCCASIONED BY CATEGORY DIVERSITY

Greek jars form a diverse category. They include vases with or



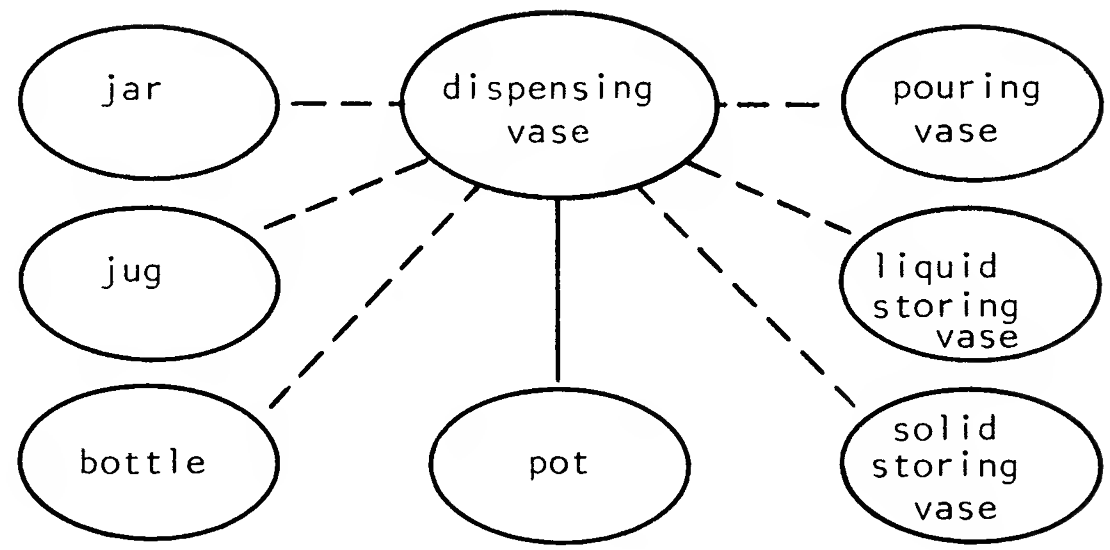


FIGURE 3.17.

FIGURE 3.16.

without necks, handles, or feet. The neck or mouth width may be broad or narrow. Different combinations of such features lead to diverse paths in the naming tree. Eventually these paths must lead via cross connections to the Greek jar module. Any of the previous three examples could be considered as cross connections forced by category diversity. A separate example is therefore not required.

3.6 Appraisal

All parts of the program except the function and name assigner work in a bottom-up manner: control passes directly from one level to the next, resulting in a description that finally leads to naming. Some interaction does occur, as between prototype selector and segmenter, when one level is unhappy with results from a lower level. That such interaction is rarely necessary is due to the domain and to the existence of a firm outline. A firm outline entered as a list of points eliminates problems of working from intensity data. The name and function assigner on the other hand is basically top down.

The segmenter has built-in assumptions about vases. If the domain were uncertain, domain characteristics would have to be divorced from the segmenter. A more top down structure might work in conjunction with the segmenter to select an appropriate domain.

The naming program could be improved by recognizing near misses. A vase may fail as an amphora only because its neck is too narrow. The present program would name the vase a jug; a more informative description might be "like an amphora, except that the neck is too narrow."

THE COMPLEXITY OF VASE DESCRIPTION

Though the program's capabilities are limited, the program does provide some index of the complexity of vase description. How to measure complexity is not at all clear; lacking anything better, I offer program size and the number of decisions as two different complexity measures.

Since these are sensitive to coding style, the numbers given are upper bounds of a sort. The program contains 3000 lines of interpretable LISP code. Counting each COND clause as one decision, there are a total of 1190 decisions: 518 to compute the descriptors and 672 to assign a name and function. There are 105 descriptive terms, listed alphabetically in table 3.4, and 53 name and function terms; the grand total is 158 terms. From the number of decisions, it appears naming and description building are about equally complex.

Table 3.4

above	globular	ovoid
abruptly	gradient	pear
articulation	gradually	pedestal
base	greatest-width	pinched-in
becoming	handles	rim
bell	height	ring
belly	hemishpere	round-bottom
below	high	segment
biconical	high-angle	shallow
body	high-shoulder	short
bottom	horizontal	shoulder
bowl	in	size
brim	inverted	slant
broad	junction	slight
calyx	large	slightly
carinated	lip	slim
circular	location	slope
concave	loop	small
cone	low	sphere
contour	low-angle	splaying
convex	low-bellied	squat
convexity	lower-profile	standard
curved	lug	stem
cylinder	middle	straight
deep	minimal	strongly
down	molded	tall
end	mouth	tallness
enormously	narrow	top
everted	neck	up
extremely	non-loop	upper-profile
flaring	not-offset	vertical
flat	offset	very
flat-base	open	whole
foot	orientation	widely
gently	out	width

CHAPTER 4 -- POLYHEDRA

Generalized cylinders model polyhedral objects with trihedral vertices (vertices formed by the intersection of three planes) particularly well. Simple constraints derived from formal considerations such as those of Huffman [1971] and Clowes [1971] lead to selection of prospective cross sections in a scene of assorted objects. By projecting such cross sections along an imaginary straight axis to form generalized cylinders, the scene is parsed into separate bodies, at the same time that descriptions are generated for them. This differs from previous work in which object separation and object identification were carried out independently. A result of the present approach is the easy handling of arbitrary alignment.

4.1 Polyhedra as Generalized Cylinders

Polyhedra are a restricted form of generalized cylinder. The axes are straight lines, the cross sections polygons, and the scale change functions linear. Possible axis positions are deduced by projecting the cross section along lines emanating from its vertices (called rays henceforth). The block in figure 4.1 can be described as the projection of rectangle A along its rays r1, r2, and r3. By connecting a point of A with the corresponding point of any projection of A along its rays, a prospective axis is determined. An axis deduction is however unnecessary because the rays suffice to guide projection and to determine cylinder length.

When a projected cross section reaches the end of one ray before the ends of all the rays are reached, the object in question is not a simple cylinder. At this point two choices are possible. (1) The object can be segmented there and the remainder described as a separate cylinder; for example, projection of cross section A in figure 4.2A could lead to a segmentation into two distinct cylinders when the ends of rays r1 and r2 are reached (figure 4.2B). (2) The projection continues to the ends of some other rays, such as r3 in figure 4.2C. Of the two decompositions, C gives the better description as block with small protrusion rather than as a smaller block with large protrusion.

The decision to continue or stop projection is therefore critical for complex object description, and must be made carefully in order to yield the "best" description. Stopping a projection always leads to additive volumes, such as protrusions or additional cylinders, while continuing a

Blank

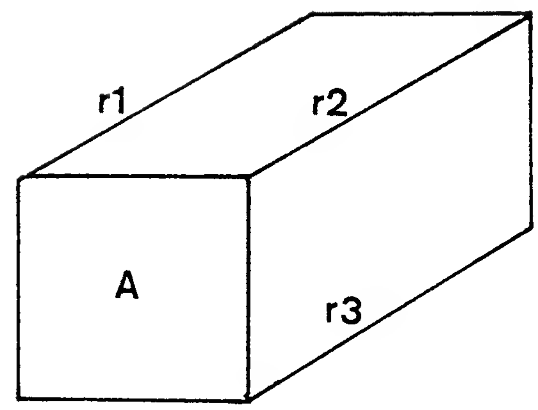


FIGURE 4.1. Cross section A projects along rays r1, r2, and r3 to form a block.

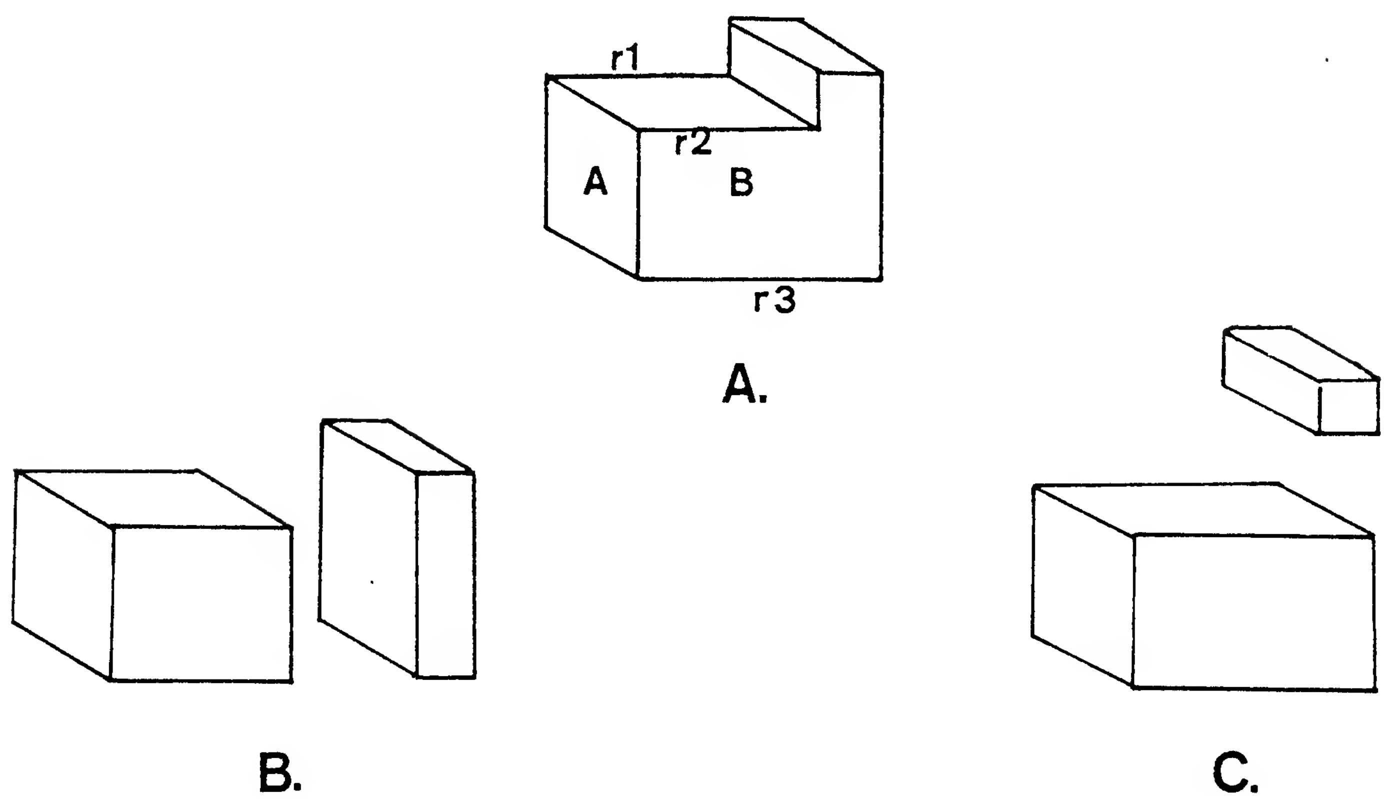


FIGURE 4.2. Projection of A can stop at the ends of r1 and r2, yielding B., or at the end of r3, yielding C.

projection may also lead to subtractive volumes, or indentations. Thus projecting A past r1 and r2 in figure 4.3 leads to a description as block with indentation.

MOST OBJECTS HAVE MORE THAN ONE POSSIBLE CROSS SECTION

The criterion for a region being a cross section is simply that its edges are convex in the 3-dimensional sense; any given object normally has several such regions. Cross sections are not allowed to have concave edges, since such a region could not encompass the whole object in the projection.

For a few objects such as cubes and blocks, the choice of cross section is largely irrelevant. For most objects, however, different cross sections often lead to vastly different descriptions. If A is chosen in figure 4.2A, the object is decomposed in one of the two ways indicated, but with cross section B the object is described as a single cylinder with an L-shaped cross section. The latter is the more economical description. One should strive therefore to choose the cross section that leads to the simplest object description, in a suitably defined sense of simplest.

Once a cross section has been selected and the projection carried out, the resulting single cylinder may require redescription. Particularly for complex cross sections, a redescription in terms of prototypes and modifiers is more suitable for comparisons. The object in figure 4.4 is well described by projection of cross section A, but A is probably too complex to catalog as a 10-sided region. Describing A instead as a

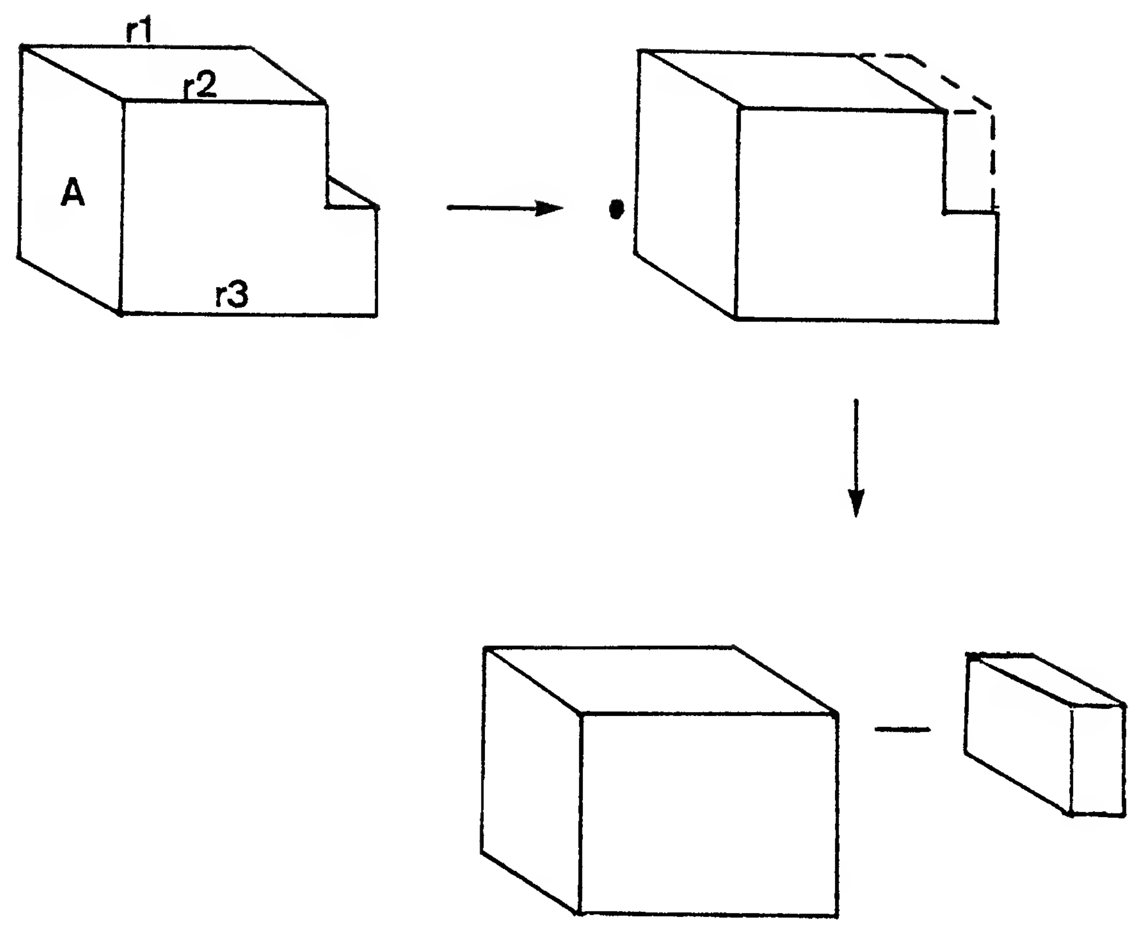


FIGURE 4.3. By projecting A to the end of r3, an indentation modification is found.

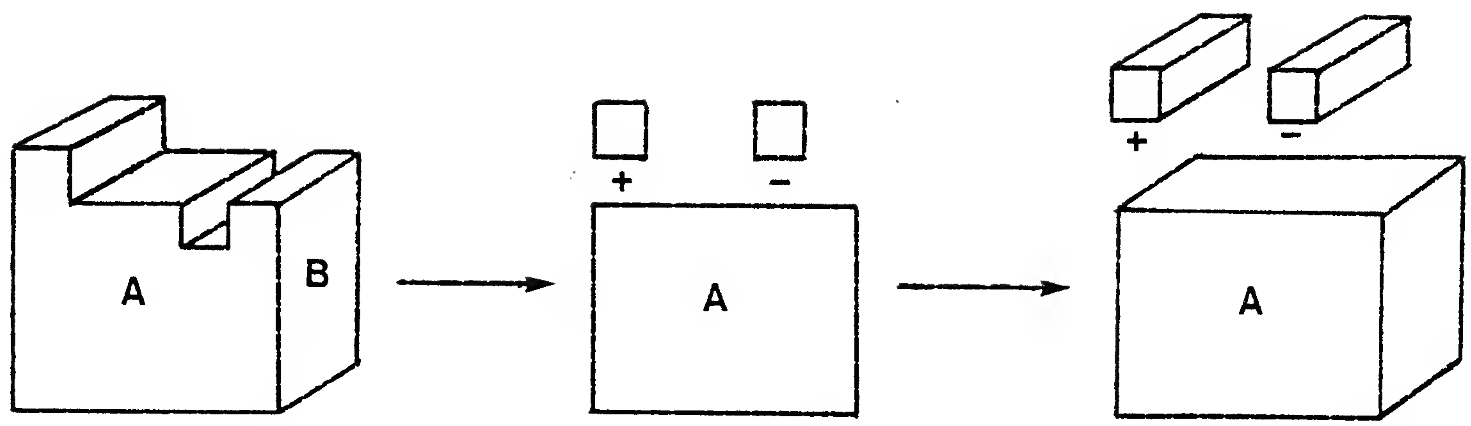


FIGURE 4.4. Decomposition of region A leads to a description as block with protrusion and indentation.

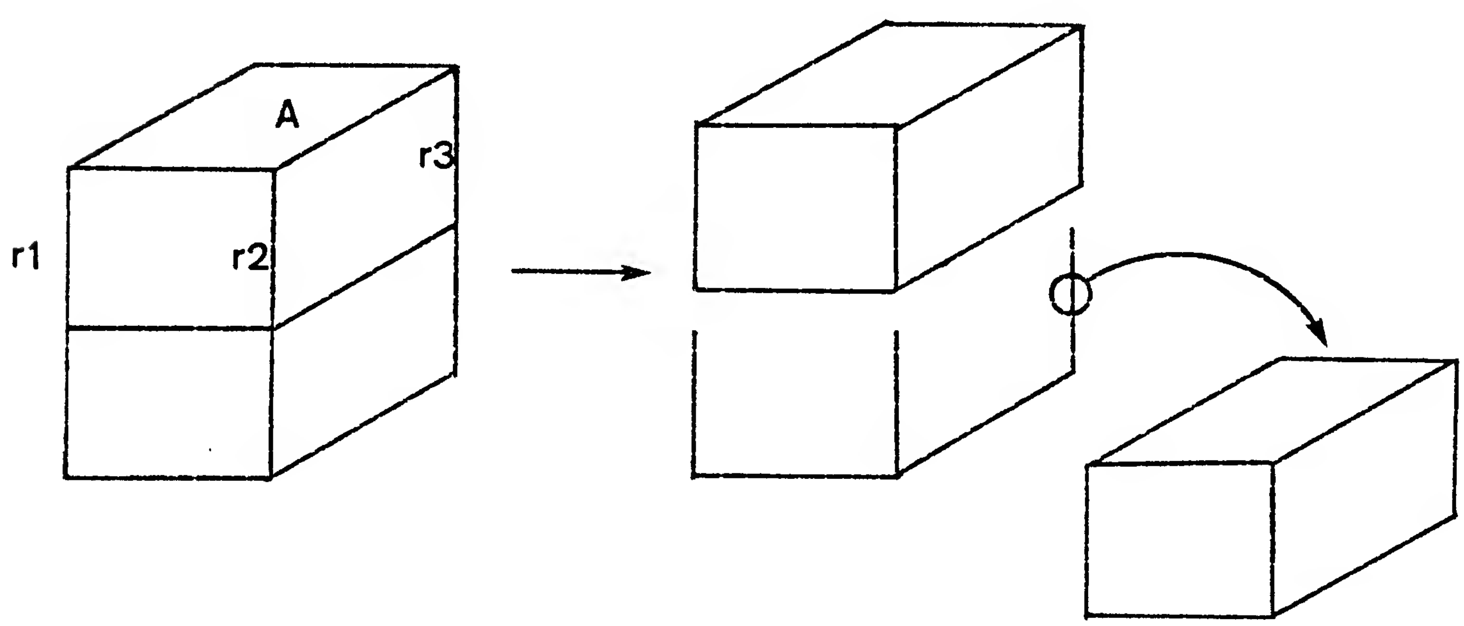


FIGURE 4.5. The bottom object is reconstructed as a block.

rectangle with a rectangular protrusion in the upper left corner and an indentation in the upper side leads to a more sensible description as a block with protrusion and indentation modifications.

There is considerable overlap between the problem of selecting prototypes and modifiers, that of choosing the appropriate termination point of a projection, and that of selecting a cross section. Choosing B, for example, and projecting it through the indentation and protrusion results in exactly the same description.

Identification of cross sections leads to a parsing of regions in a scene into bodies, because the regions associated with the rays of a cross section are grouped with it. Thus projection of A in figure 4.5 along its rays leads to the identification of the top block.

Analysis of the scene can continue with a deletion of recognized objects from the scene so as to unobscure others. Subsequently, the character of the obscured portions must be guessed at; for example the bottom object in figure 4.5 should probably be reconstructed as a block.

WHAT DETERMINES A POSSIBLE CROSS SECTION?

A region is a possible cross section if there are a set of lines that can be interpreted as rays in a consistent manner. This is a precise way of stating what it means to look like an object in the polyhedral domain. Surprisingly selection of rays for a cross section can be done in a simple, automatic manner with rules that can be presented in the form of a finite state machine, presented in the next section.

In conclusion, the application of the proposed description methodology to scenes of polyhedra leads to the following steps:

1. Selection of prospective cross sections.
2. Deletion of recognized objects and reconstruction of obscured objects that become unobscured.
3. Choosing the best cross section for description once an object has been separated.
4. Determination of the termination point for a cross section projection.
5. Description of a cross section in terms of prototype and modifiers.

These steps are not completely independent, and can interact in complex ways. The considerations that apply at one step, moreover, may also be necessary for another. The remainder of section 4 deals with these steps in more depth.

4.2 Cross Section Selection

The restriction of cross sections to regions with convex edges means that the edges of such a region in a two dimension projection can have only convex (+) or obscuring (<-) line labels (an obscuring edge is a convex edge that has only one face visible). More formally, a cross section may have only type 1 and type 3 vertices [Huffman 1971], which are listed in figure 4.6A.

Huffman types vertices by examining how many octants are filled with solid material with the vertex as origin. A type 1 vertex corresponds to any way a vertex can be viewed from the complementary 7 octants when one octant is filled. A type 3 vertex correspondingly fills 3 octants. Concave edges are indicated by a '-' labeling. Any region whose vertices are a combination of these 7 vertices can be a cross section, with the obvious constraint that a line from one vertex match the corresponding line label of the vertex to which it is connected.

The lines of a scene are not prelabeled of course, and to identify cross sections it is necessary to work in the other direction: how must the vertices of a region look so that they could be interpreted as type 1 or 3 connected in a permitted fashion? To aid in the subsequent discussion, Waltz's region labeling in figure 4.6B will be extended. Waltz's region labels refer to particular regions partitioned by a vertex type, such as the A1 region of the arrow vertex. I will also speak of A1 as a vertex type: an A1 vertex is an arrow whose A1 region coincides with the region in question.

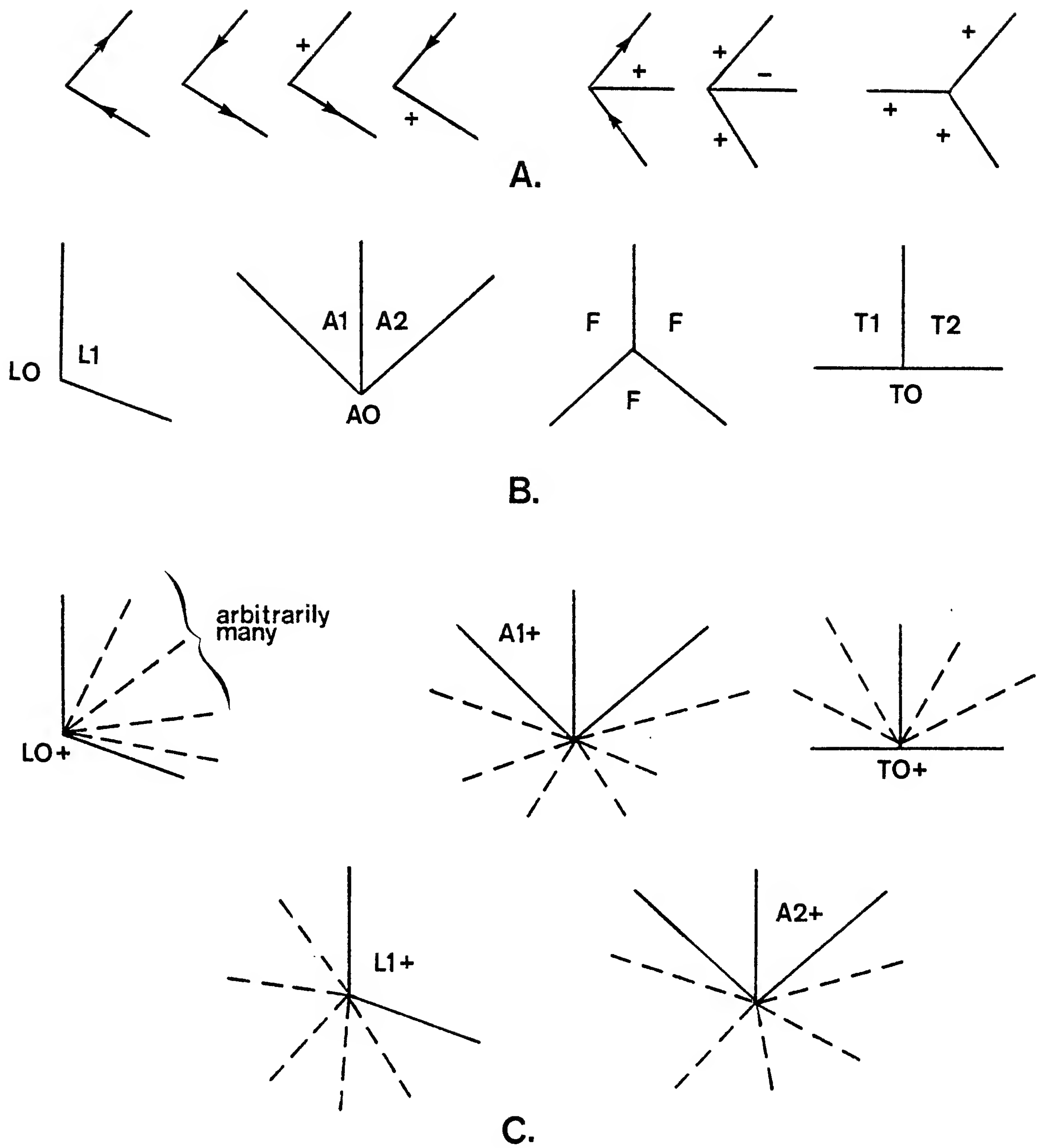


FIGURE 4.6. Some notation.

USING HUFFMAN LABELING TO DERIVE CONSTRAINTS ON CROSS SECTIONS

The following discussion assumes isolated bodies. Alignment will be dealt with later.

Vertex combinations in figure 4.6A are restricted because of the necessity of a common line label. For example, an L1 vertex of a cross section can only receive the first L labeling because the remaining 3 are characteristic of L0 vertices. Moving in a clockwise direction, the obscuring edge of the L1 vertex can only attach to the obscuring edge of the first arrow, the fourth L, or the first L again. Thus an L1 may be followed by an A2, L0, or L1 vertex. Similarly, a cross section with an A2 vertex may only connect to an A1, L1, L0 or F vertex.

By carrying out this process for all the vertices, a transition net is obtained that concisely summarizes these restrictions. The transition net can be represented in a number of equivalent ways, such as a linear grammar, or as the finite state machine in figure 4.7. In the remainder of this chapter, the FSM representation will be used.

The transition net was also derived in part by Waltz [1971] as a regular grammar for type 1 vertices around a region. The present formulation goes considerably beyond Waltz's original grammar insofar as it also includes type 3 vertices and has been extended to handle alignment.

The FSM can be used to recognize cross sections. It accepts a region as cross section if it starts in any state and returns to that state so that (1) at least three states are entered (not necessarily distinct), and (2) at least one of the states is from the set [A0,A1,A2,F]. The first

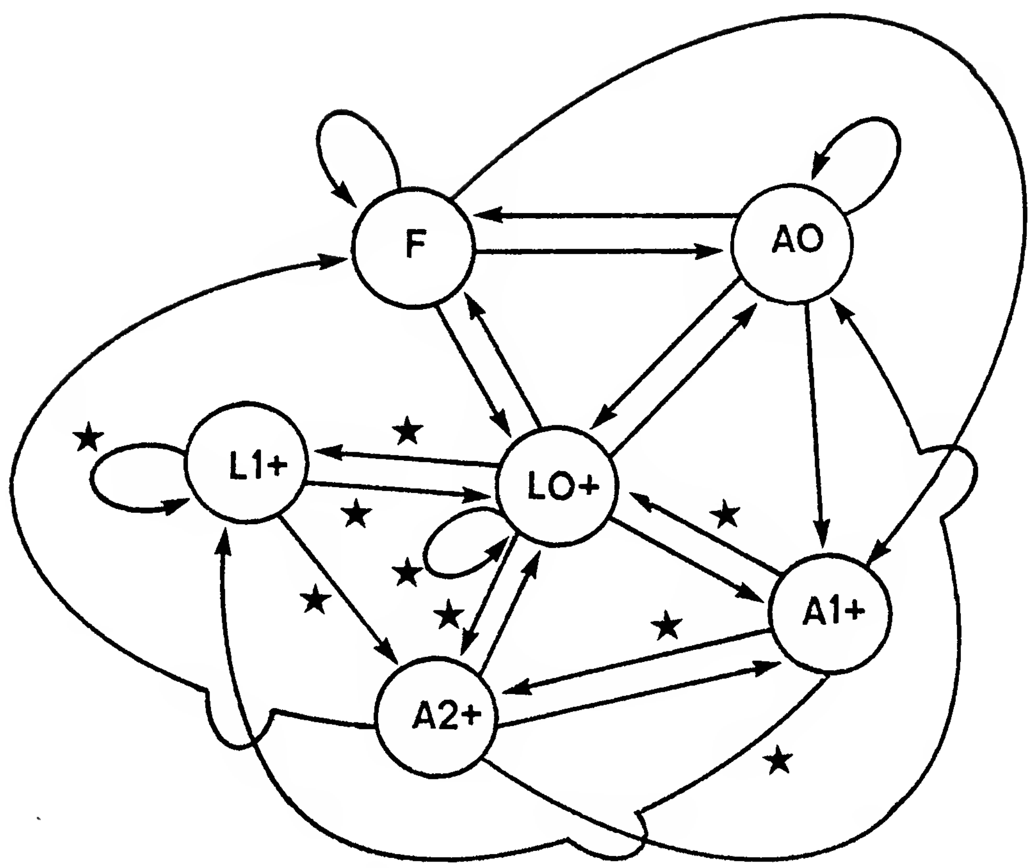


FIGURE 4.7. FSM for scene parsing.

condition is an obvious one requiring a region to have at least three vertices, but the second requires further discussion for justification. The '*' and '+' marks will be explained later.

CROSS SECTION CONSTRAINTS CAN ALSO BE DERIVED BY EXAMINING RAY TOPOLOGY

A more intuitive way of deriving the constraints, a way that also gives more insight into what the FSM is doing, is to examine topological restrictions induced by a given vertex on rays of neighboring vertices. Suppose there is a convex region angle whose ray makes a convex angle with its clockwise side (i.e., an A2 state as shown in figure 4.8A). During projection along the ray, edge e1 of the region remains parallel to its original position. This is a result of the existence of a straight axis. Vertex v, the other end of e1, describes a line v-v' during projection (figure 4.8B). This line need not be parallel to the ray, but it must be straight because scale change is linear. It must also lie on the same side of e1 as the ray. There are two cases to consider: (1) the region angle at v is convex, and (2) it is concave.

(1) The region angle is convex. Then the ray v-v' must be visible. If it were obscured, it would lie on the opposite side of e1 as the first ray, violating an earlier observation. The topological alignment of the ray v-v' and the next edge e2 of the region may yield either an arrow (figure 4.8C) or fork (figure 4.8D) vertex type for v. That is to say, an A2 vertex may be followed by an A1 or F vertex.

(2) The region angle is concave. This time the next edge e2 of the

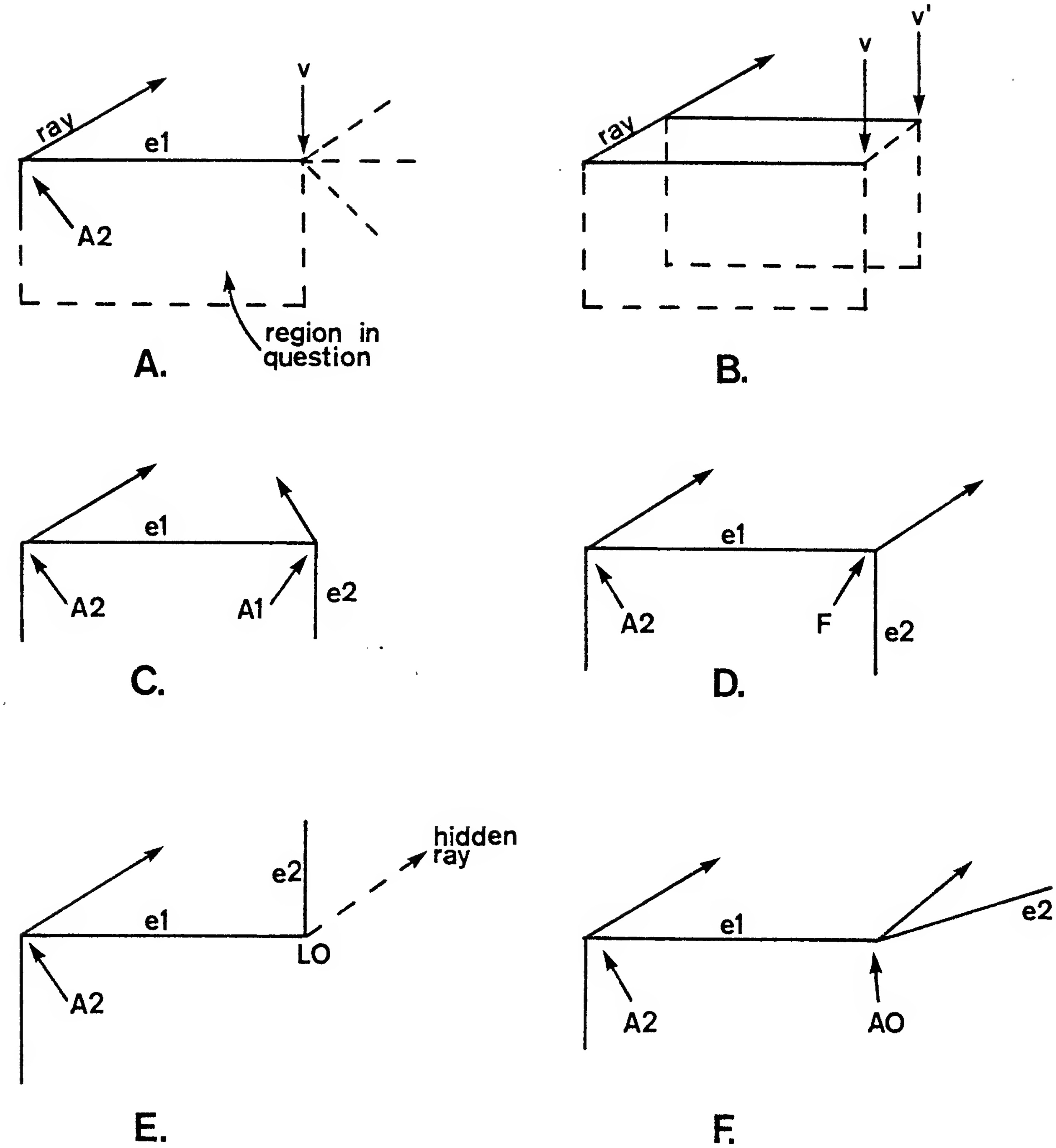

ray
e1
v
A2
region in question
A.
ray
v
v'
B.
e1
A2
A1
e2
C.
e1
A2
F
e2
D.
e2
hidden ray
e1
LO
A2
E.
e1
e2
A2
AO
F.

FIGURE 4.8.

region may obscure the ray v-v' (figure 4.8E). When it does not, the situation in figure 4.8F is derived. Thus an A2 vertex can also be followed by an L0 or A0 vertex.

Note that these four possibilities are the only ones giving rise to transitions from A2 to another state. This process can be carried out for all vertex types, yielding the transition net in an alternate manner.

EXPANDING THE CONSTRAINTS TO HANDLE ALIGNMENT

Alignment complicates cross section recognition either by camouflaging rays in a thicket of non-region lines, or by obstructing a ray entirely. Region A of the wedge in figure 4.9 illustrates the first complication. It is aligned with the bottom block, and at vertex v either e1 or e2 (or neither) could be the ray. The "neither" case is illustrated by the L1 vertex w of region A, which has 3 non-region lines but not a ray.

The non-region lines e1, e2, and e3 do not interfere with projection of A because the region would move away from them. Such non-interfering lines of alignment are indicated by a '+' next to a vertex type; for example, w would be labeled L1+. Similarly, vertex v is written as A1+, where e1 is the ray and e2 is assigned to the '+' category.

This alteration applies to all vertex types except F and A0. For the latter two, an edge of the projected region moves along either side of the ray. If another non-region line were present, it would act as an obstruction to one of the projecting edges.

A T0 vertex also represents a form of alignment, and may occur only

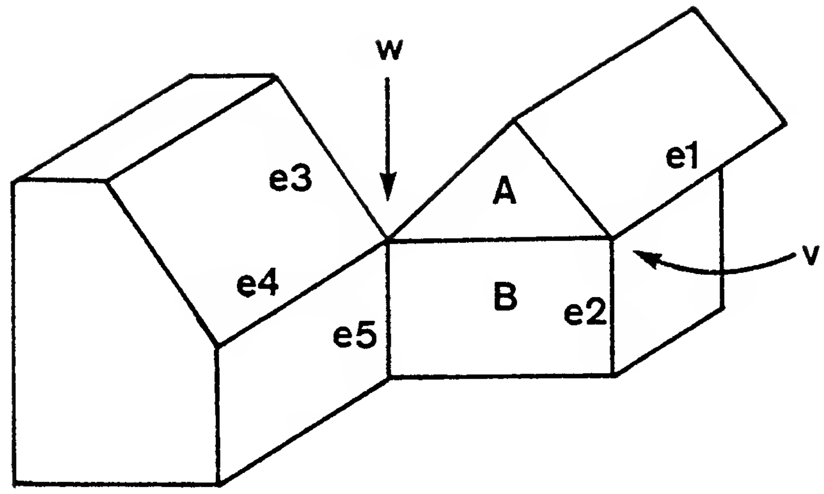


FIGURE 4.9.

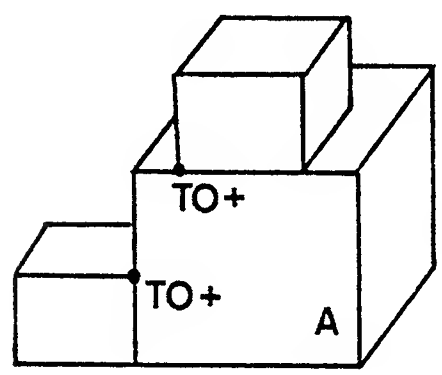


FIGURE 4.10. The left TO+ is allowed by the FSM, but the top one is not.

when the projection is away from it. More precisely, it cannot appear after a visible ray vertex where the shaft of the T0 is on the same side of the connecting edge as the ray. Thus the T0+ vertex on the left side of region A in figure 4.10 does not interfere with projection of A, whereas the one on the upper side does. This restriction is indicated by '*' marks on some transition arcs of the FSM, meaning that during such transitions an arbitrary number of T0+ vertices may appear along the connecting edge.

WITH THIS EXPANDED LABELING, VERTEX TYPING CAN BE AMBIGUOUS

The source of ambiguity is the '+' category. A fork vertex, for example, may be interpreted as F or L1+. Vertex w in figure 4.9 could be assigned L1+, A2+, or A1+ with respect to region A. Vertex ambiguity is resolved by finding a consistent assignment of non-region lines into the '+' or the ray category for all vertices of a region. When successful, the region is projectable along the discovered rays without interference from the other non-region lines. Otherwise, another region must be chosen as cross section.

Note that the restriction of the FSM to enter at least one of the states that predicts a ray, i.e., one of [A1+,A2+,A0,F], rules out the possibility of interpreting all vertices as L1+ or L0+. This assignment is in principle possible for every region, but it is trivial and hence is ruled out. Guzman's [1968] proof about the realizability of any scene is an equivalent observation.

Besides this trivial assignment, a region may legitimately have more

than one assignment if there is an ambiguous scene where a cross section may be projected in two different ways. The familiar example is figure 4.11, which may be decomposed in the two ways shown. The first results from projecting A along rays e1 and e2, the second from projecting along e3 and e4. Both interpretations are found by the FSM.

The assignment of vertex types could be made more efficient by starting with the less ambiguous ones. Completely unambiguous is an L vertex without '+' edges. Subsequent assignments might move in a clockwise direction from the L vertices. If there are no L's, it is probably best to work from vertices with only one non-region line, etc. Alternatively, each vertex can be assigned a list of all possible interpretations. Restrictions between neighboring vertices can propagate around the region in a Waltz-like manner until a consistent assignment (or two) is found, or all lists are depleted. I am indebted to Gene Freuder for the latter suggestion.

CROSS SECTION SELECTION CAN LEAD TO UNREALIZABLE OBJECTS

Cross section recognition is a local process, encompassing only a region and its rays. What happens at the other end of the rays is not taken into account, and may invalidate the existence of the purported object. Because of the curious alignment of blocks in figure 4.12A, region A appears projectable along rays L1, L2, and L3. One of the regions encompassed in the projection, unfortunately, corresponds to the background as evidenced by a missing line between L1 and L2. Another example is

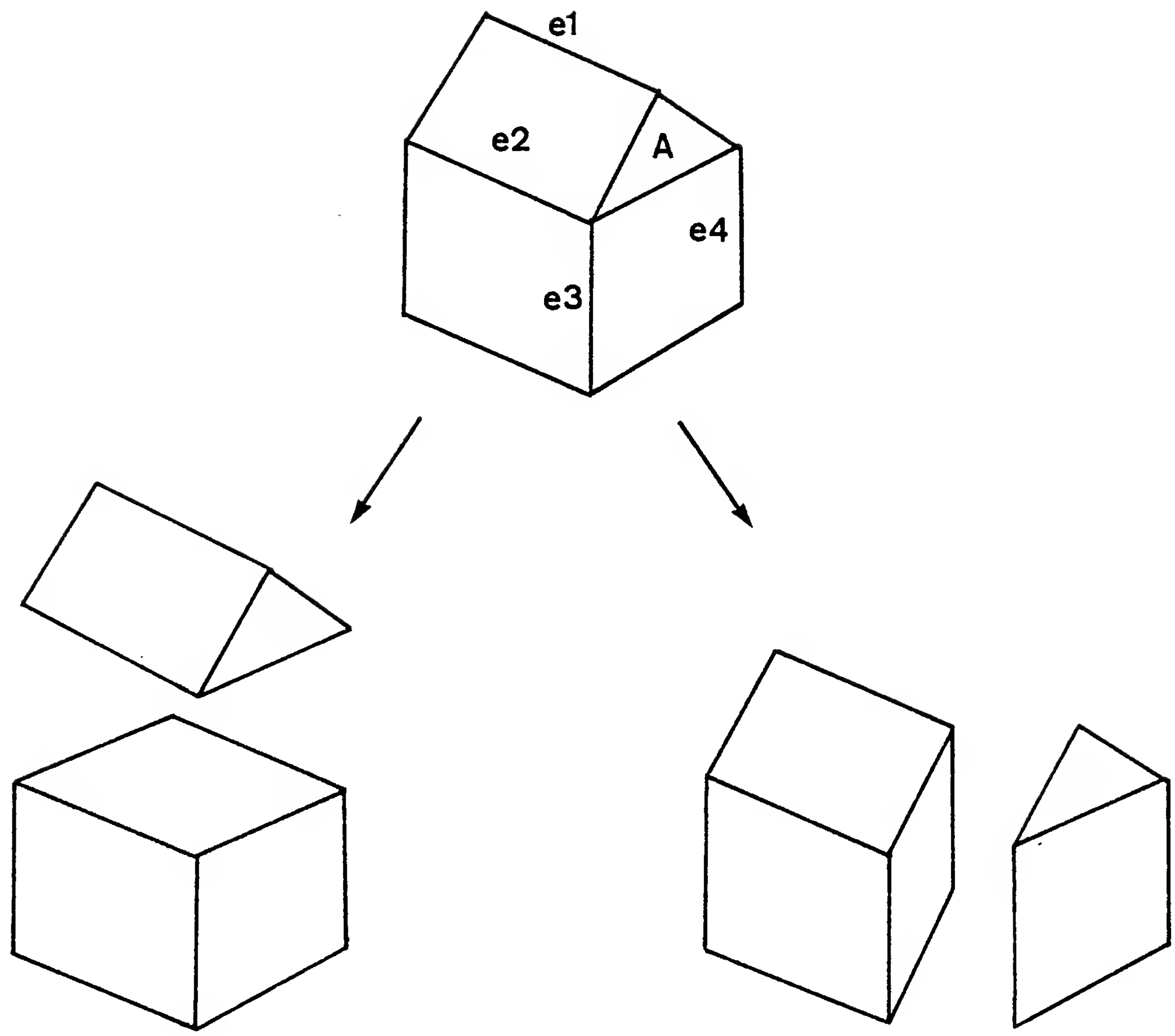


FIGURE 4.11. Depending along which lines A is projected, different decompositions are obtained.

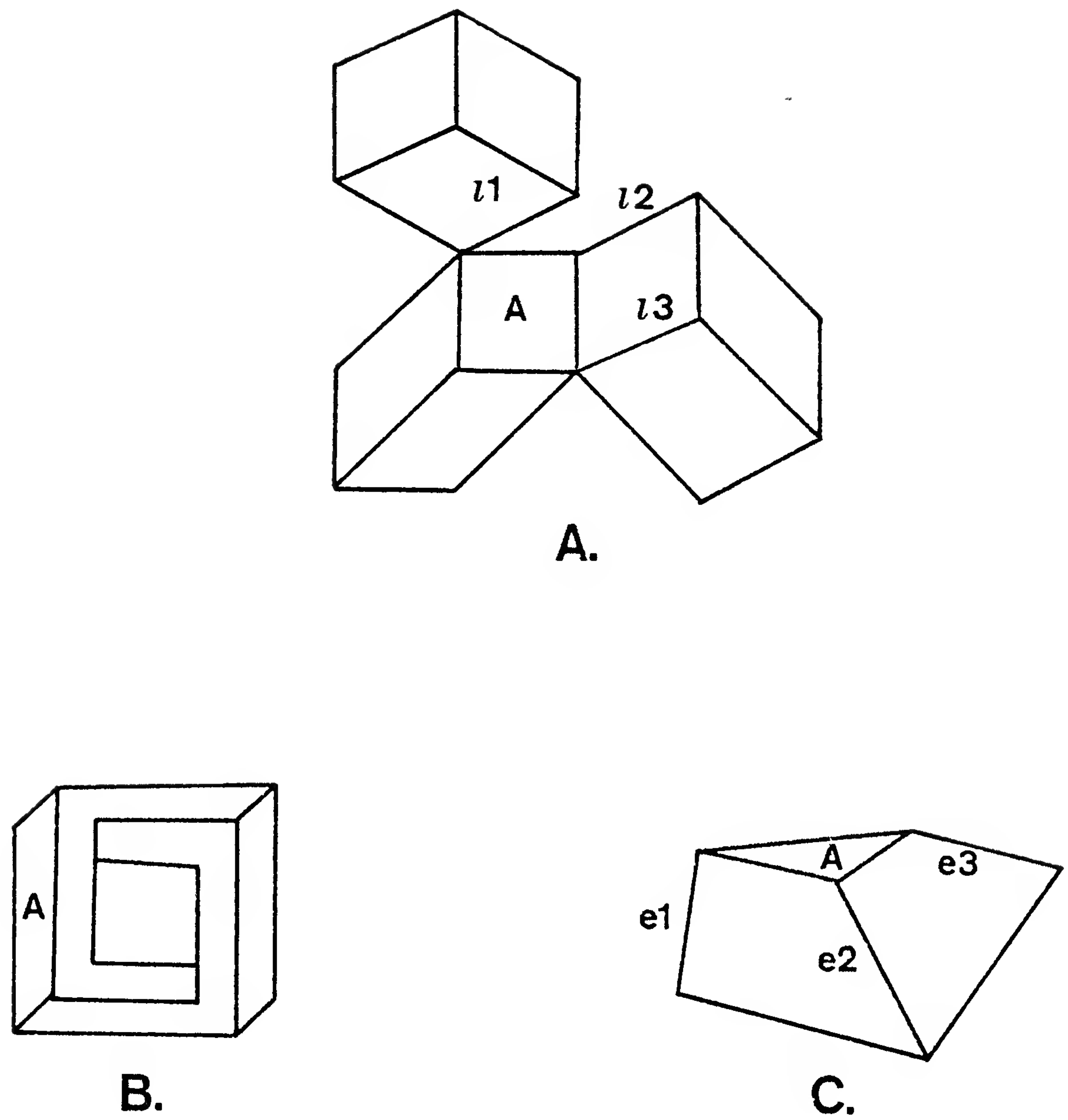


FIGURE 4.12.

region A in figure 4.12B, which though projectable soon runs into irreconcilable conflict because the object is nonsensical.

Both examples appear solid anyway because A can be projected a short distance before a problem is encountered. An optical illusion resulted when the apparently valid projection was interrupted in an unexpected way. This is almost a formula for creating a class of optical illusions: create an irreconcilable obstruction to the path of a projection.

Another source of unrealizable objects recognized by the FSM arises from the exact positioning of rays. When rays are not positioned in a strict quantitative relationship, the object cannot be physically realized with only trihedral vertices. If snapshots of a cross section were taken as it was projected along its axis, the cross section at different intervals would have the same orientation and remain geometrically similar except for a possible scale change factor. It would have the same orientation because the axis is straight and the cross section is constrained to maintain the same solid angle with respect to it. It is geometrically similar because the scale change function results in a proportionate change in length of the sides. In the case where scale change is not zero, the cross section must eventually collapse to a point when hypothetically projected far enough along the axis in the appropriate direction. Thus the rays e1, e2, and e3 of cross section A in figure 4.12C must meet in a single point when extended, yet they do not.

Between any two snapshots, it can be deduced from this discussion that corresponding sides of the cross sections are parallel. This is another

way of looking at Huffman's unity gain criterion for the realizability of trihedral polyhedra. The FSM is therefore too lax about requirements it puts on rays. This laxness is not serious, however, because the object looks real all the same. Many of us would actually have to apply Huffman's criterion to be convinced otherwise.

To summarize the last few paragraphs, the FSM determines which regions might lead to formation of a body, but only the process of projection itself can indicate if the resulting object is physically realizable. What it finds, however, will look at least partly real.

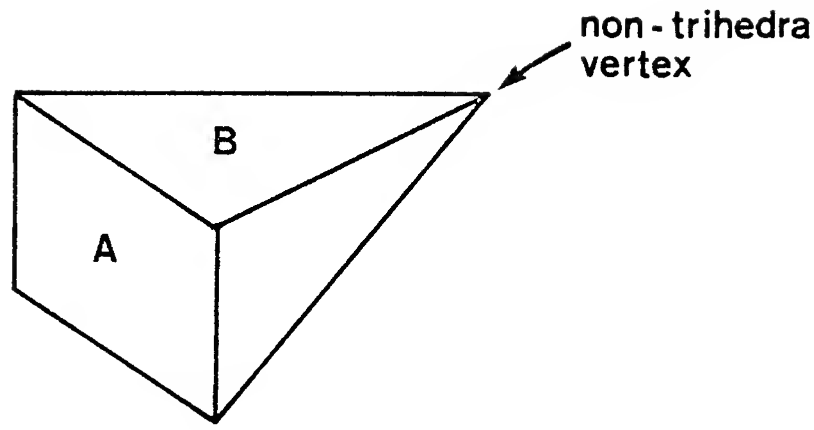
Perspective deformation has not been taken into account, and changes how objects appear. Cross section recognition is not affected by perspective deformation, but the actual process of projection must be modified to take foreshortening into account. The side lengths of projected cross sections, in particular, will not be observed to maintain the same ratio.

SOME NON-TRIHEDRAL OBJECTS ARE ALSO RECOGNIZED

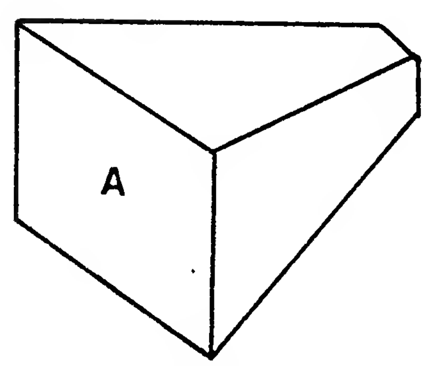
There is a close relationship between some trihedral and non-trihedral polyhedra. Any cross section with a linear scale change function will, if allowed, project to a point. This point forms a non-trihedral vertex (except for triangular cross sections), with as many edges as there are sides of the cross section. When the projection stops short of a point, the object is trihedral.

The FSM is able to recognize this type of non-trihedral object, as

illustrated by objects A and B of figure 4.13. Region A would be successfully recognized as cross section in both objects, and would project correctly for object A even though non-trihedral. On the other hand, region B does not work as a cross section because its scale change function is not uniform for all sides. Projecting B does not reveal the existence of the hidden line at the non-trihedral vertex; rather, it is indicated by projecting A.



A.



B.

FIGURE 4.13. Non-trihedral vertices are handled successfully if the projection ends there.

4.3 Scene Parsing

In a scene where some objects partially obstruct others, the obstructed objects cannot usually be recognized because potential cross sections have hidden rays or are obstructed in the path of projection. Scene parsing must therefore proceed by "unstacking": cross sections of unobstructed objects are found first, and the objects formed from them are deleted or "unstacked" from the scene. Thus the obstruction of the remaining objects is reduced.

By deleting such objects, previously hidden parts come into view. For scene analysis to continue as before, the nature of these hidden parts must be conjectured. Guzman [1968] and Waltz [1972] advocated scene parsing without knowledge of identity as positive features of their respective approaches. Conjecturing hidden parts is not in question here, for it must be made sometime for identification; rather, the question is at what stage it takes place. The present approach suggests that picking out an object and identifying it go hand in hand.

Context and real world constraints participate in selecting the order of examining regions and in conjecturing hidden parts. The simple procedures presented below for this purpose should be seen as one way of applying such knowledge, rather than as a definitive way of conducting the analysis. Though too simplistic and general to be completely adequate, they work reasonably well.

PROJECTIONS MAY TERMINATE IN A NUMBER OF WAYS

For unobstructed objects, one of three situations results when a projection terminates:

1. All regions of an object are encompassed in the projection.
   This happens when each visible region is formed by a projecting edge of the cross section.

2. Not all regions are encompassed, but there is a better choice that does encompass all regions.
   For example, cross section A in figure 4.14A does not encompass regions C and D; E should have been chosen because it encompasses them all.

3. No cross section choice encompasses all regions.
   None of the potential cross sections A, B, or C in figure 4.14B encompasses the three regions D, E, and F. However, by projecting each one separately to bind regions, all the regions are effectively linked to one body.

The problem of selecting which of several possible cross sections to represent an object or how to segment complex objects into single cylinders is discussed later.

For partially obstructed objects, the path of projection may be blocked; for example, A in figure 4.14C cannot finish its projection because of the other block. There are two choices at this point: project past the obstruction to a natural termination point; or (2) remove the obstructing object before proceeding. For mutually obscuring objects, the former choice is the more reasonable one. This is a special case that could be dealt with by special means when it arises. Hence it is disallowed in the remaining discussion, so that if the path of a projection is blocked, the obstruction is removed.

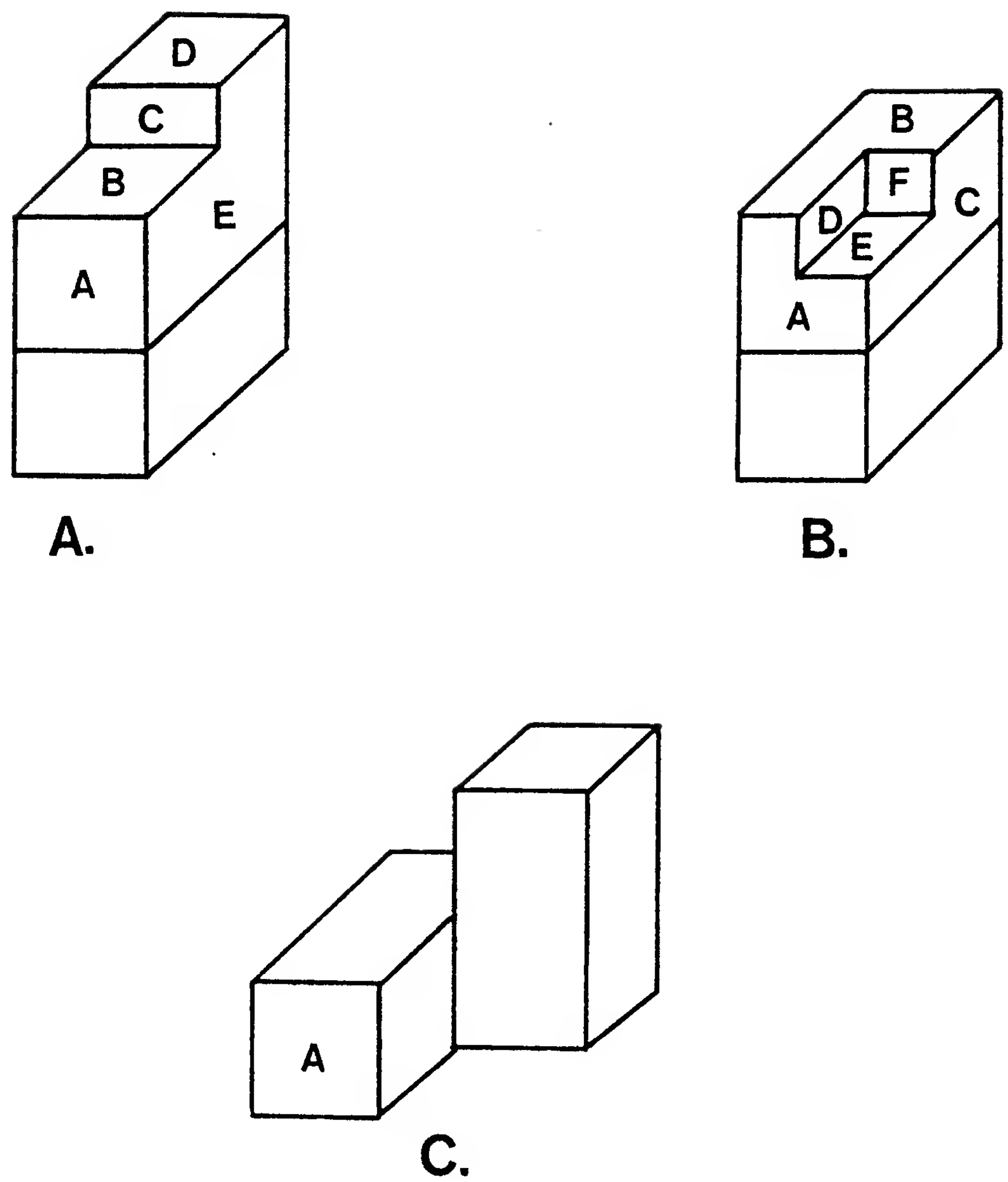


FIGURE 4.14.

AFTER AN OBJECT IS DELETED, THE SCENE IS RECONSTRUCTED

Once an object's regions have been linked, it is deleted from the scene by erasing the associated lines. Edges of other objects aligned with the object edge may also be deleted as a result, but this is unavoidable because of a lack of prior knowledge of alignment. Such edges are reinstated during the reconstruction phase during which previously hidden parts are conjectured.

Five simple rules, arranged below from more to less certainty, do a fair job of reconstruction, and were derived from constraints and likelihoods of the domain.

1. Join a split edge (figure 4.15A).
2. Extend two lines to a corner when this makes sense (figure 4.15B).
3. Extend parallel lines between neighboring regions (figure 4.15C).
4. Hypothesize a best completion when lines are parallel or do not meet at a reasonable spot (figure 4.15D).
5. Complete a region as a parallelogram when only two connected edges are present (figure 4.15E).

The first three rules are easy to understand. The fourth restores common edges erased during object deletion. The last constructs a totally obscured region in the simplest way possible, as a four-sided parallelogram. These are the most prevalent of regions because they play the role of 3-D "filler": they form the sides of projected cross sections.

Such simple rules are inadequate in situations where contextual knowledge is needed to mediate a reconstruction. Thus the bottom object in

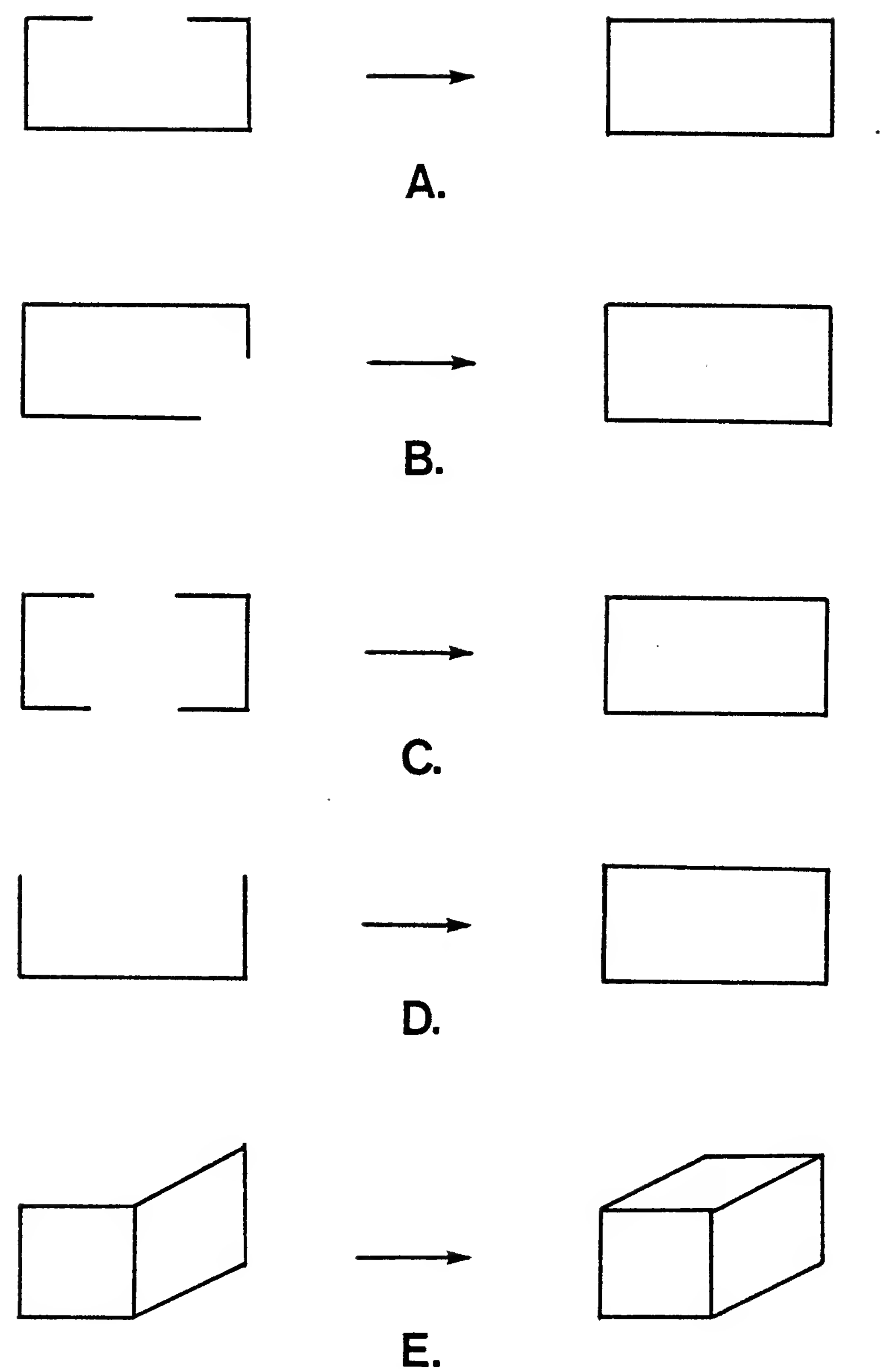


FIGURE 4.15. Reconstruction Rules.

figure 4.16A should because of context be interpreted as a wedge, but rule 5 would reconstruct it as a block.

Finin [1972] has applied some forms of high-level knowledge to aid in such conjectures. He uses the context of the top wedge to predict the bottom wedge. He also uses real-world constraints to set bounds on the dimensions of partially obscured objects. For example, the top block in figure 4.16B has uncertain length because of its uncertain distance from the bottom one; therefore Finin's program sets bounds by determining its minimum and maximum possible distance from the bottom block.

WHEN A REGION FAILS TO QUALIFY AS CROSS SECTION, THE REASON FOR FAILURE CAN LEAD TO A BETTER CANDIDATE

A simple procedure for recognizing objects is to find all cross sections, fashion objects from them, delete these objects, reconstruct the scene, and repeat the process. A more knowledgeable approach might use the results of a failure to recognize a region as cross section to suggest which region to try next. Failure often results from partial obstruction of a region by another object. The latter object is likely less obstructed; hence attention should be transferred to one of its regions. A chain of such failures is followed until an unobstructed object is found.

Finding obstructions hinges on finding the vertex at which the FSM cannot make an assignment. One way this can happen is as follows:

<u>Failure condition 1:</u> A forbidden T0+ vertex is encountered.
T0+ vertices cannot follow F, A0, or A2+ vertices. When they do, another object is aligned with the connecting edge. A region of this

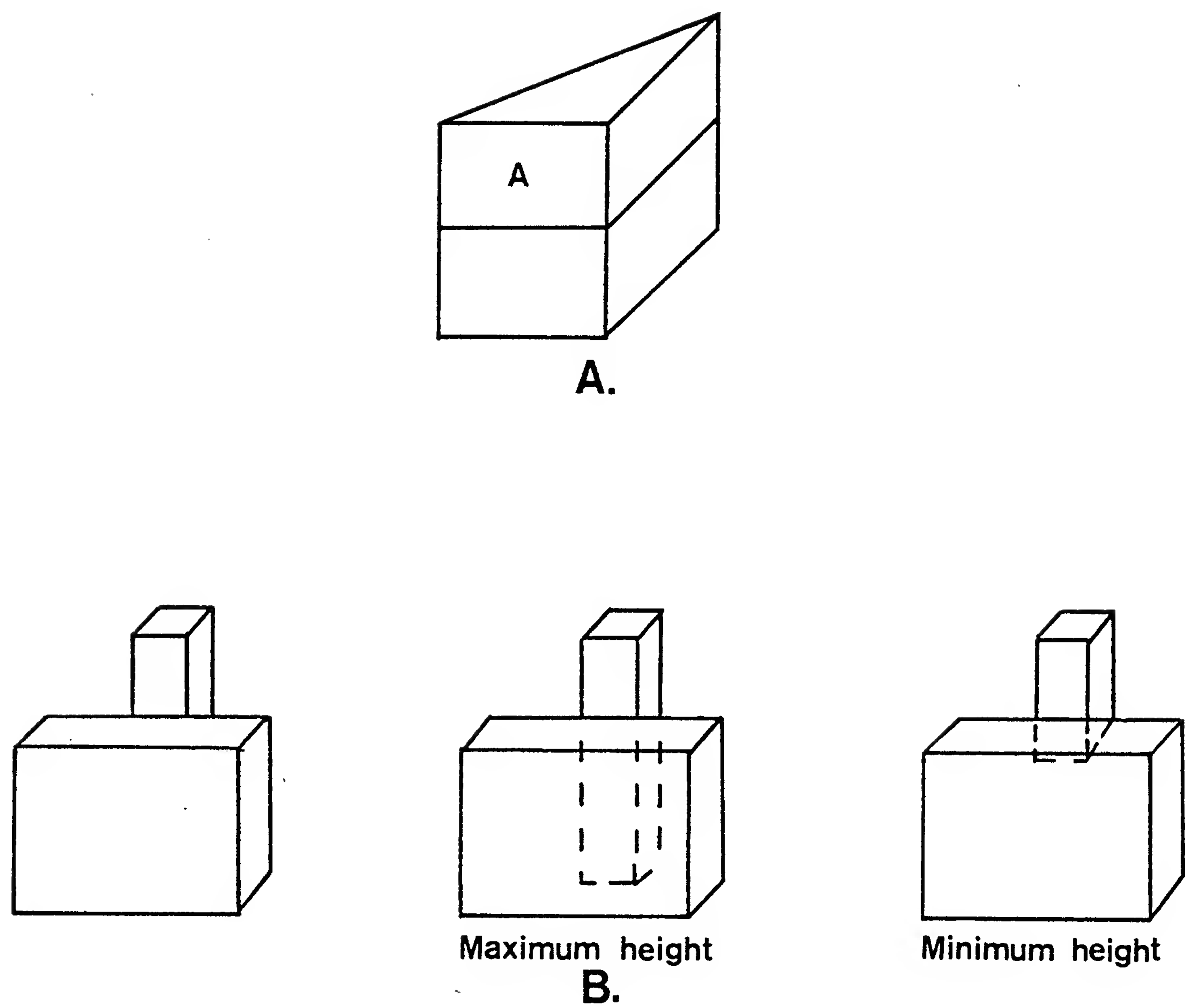


FIGURE 4.16. High level knowledge is needed to complete some partially obscured objects. These examples are taken from Finin (1972).

> obstructing object can often be located between the shaft of the T and the clockwise portion of the connecting edge. If there is more than one "shaft", look at the region between the first two shafts.

More failure conditions could be added, but an illustrative parsing will be presented below with just this one.

A SCENE IS PARSED TO ILLUSTRATE RECONSTRUCTION AND THE USES OF FAILURE

Suppose the analysis begins with region 1 in figure 4.17A; this region is apparently the largest, and its choice can be justified on this basis. The lower left vertex is the only unequivocal one, receiving an L1+ label, and the FSM will proceed clockwise from it. The next vertex is either an L1+ or A2+; however, L1+ leads to L1+ all the way around, which is a disallowed interpretation. Hence A2+ is left by default. A2+ immediately causes a snag because of the T0+ at the junction of regions 1, 3, and 5. Failure condition 1 applies here, and shifts attention to region 5.

The lower left vertex of region 5 is forced to be an L1+, and as before the next vertex is either an L1+ or A2+. The L1+ once more leads to L1+ all around, forcing an A2+ assignment. A T0+ vertex at the junction of 5, 6, and 7 causes a snag this time, and failure condition 1 shifts the focus to region 7. At least region 7 is accepted as cross section, and forms a block with regions 8 and 9.

The next step is to delete the object (figure 4.17B) and to reconstruct the scene (figure 4.17C); rule 1 accomplished the reconstruction. Now region 5 is projectable, and deleting the resultant wedge leaves figure 4.17D. Rule 1 completes region 1, while rule 4

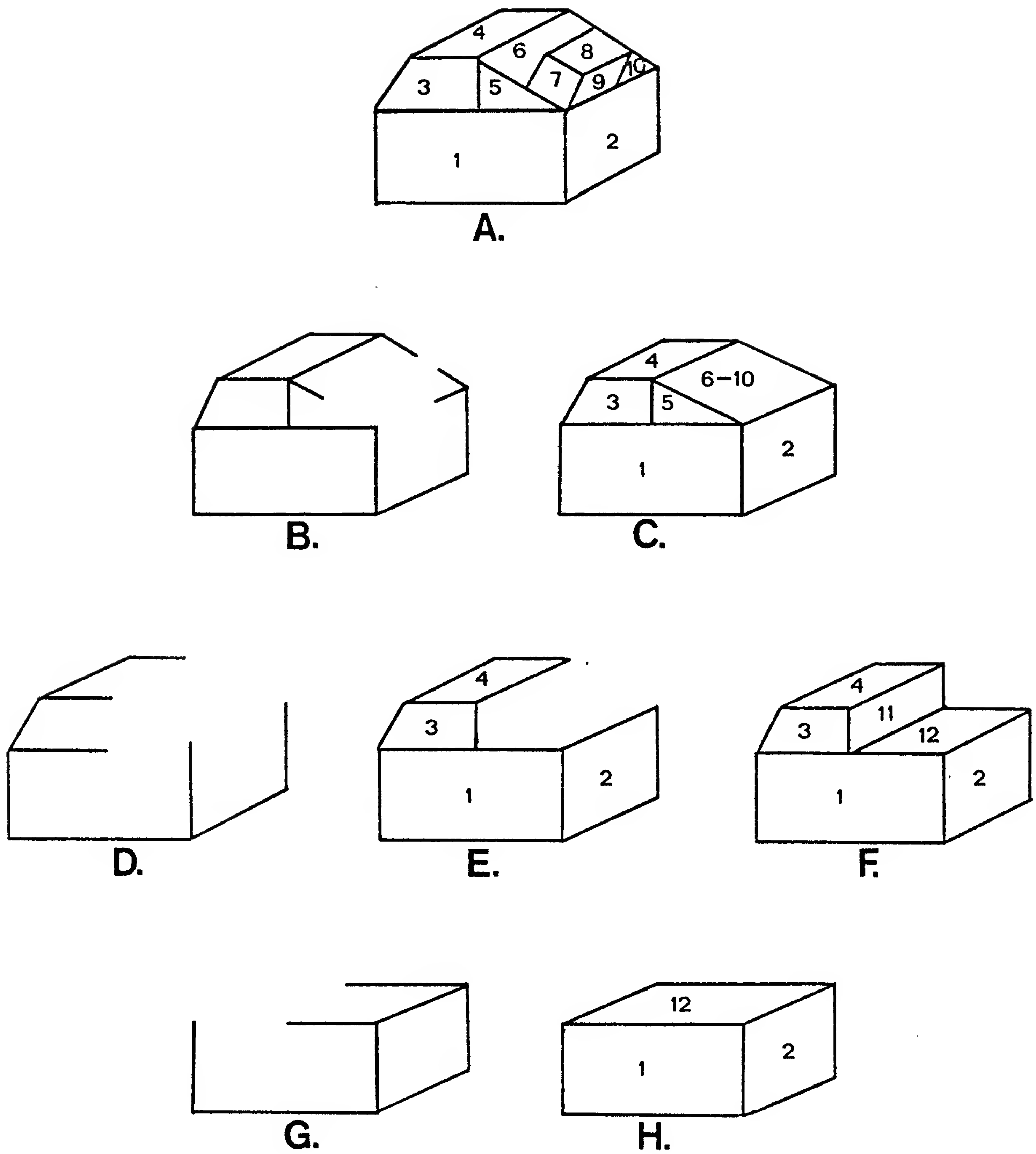


FIGURE 4.17. Papa.

reconstructs regions 2, 3, and 4 (figure 4.17E). The parallelogram hypothesizer, rule 5, completes the reconstruction by postulating two new regions (figure 4.17F).

Even though two objects have now been deleted, region 1 still fails as cross section. The culprit is the newly constructed T0+ at the junction of regions 1, 3, 11 and 12. The failure condition this time pinpoints 11, which can be projected to form a block with regions 3 and 4 (in the descriptive phase, region 3 should replace 11 as cross section to yield the simpler description as a trapezoidal block). Object deletion (figure 4.17G) and scene reconstruction by rules 1 and 4 (figure 4.17H) now allows cross section 1 to form a block with regions 2 and 12.

DIFFERENT INITIAL REGIONS MAY YIELD DIFFERENT SCENE PARSINGS

The interpretation of a scene may depend on which region is examined first. Most people see figure 4.18A as three stacked blocks and a wedge. The scene is ambiguous, however, since region B can be projected to yield the object in figure 4.18B. Why do people see the first interpretation, but not the second unless it is pointed out to them? One might argue that people work inward from regions bordering against the background; this is a contour-based approach as advocated in section 2.3.1. Other possible reasons against the second are considerations of gravity, support, and general position.

By applying the appropriate regions to the FSM, every possible scene parsing could be found. If one were interested in only the "conventional"

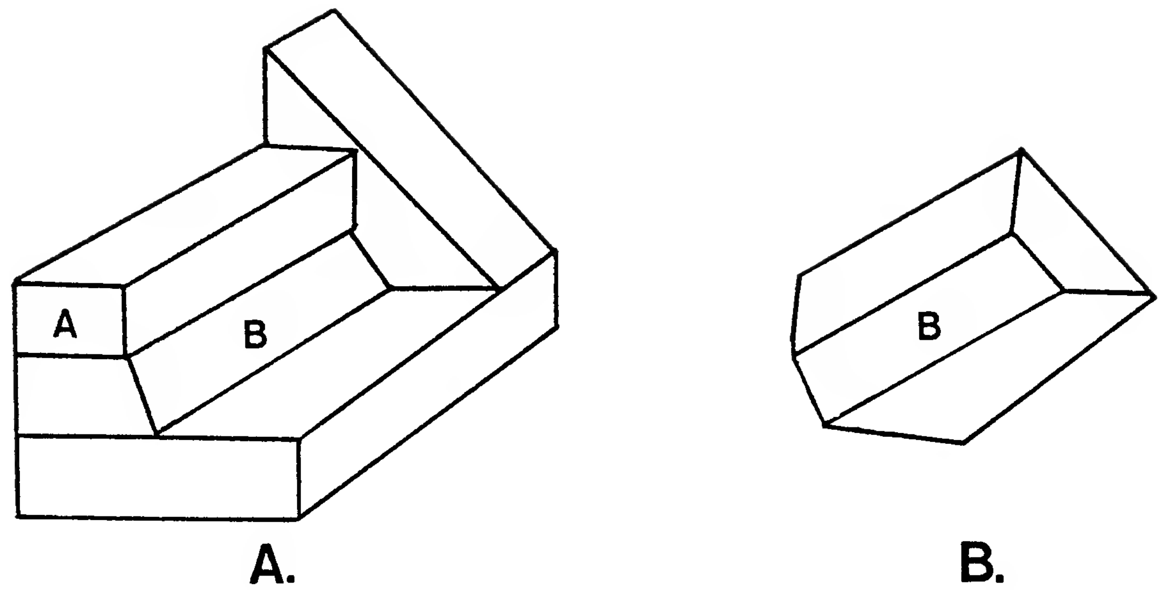


FIGURE 4.18.

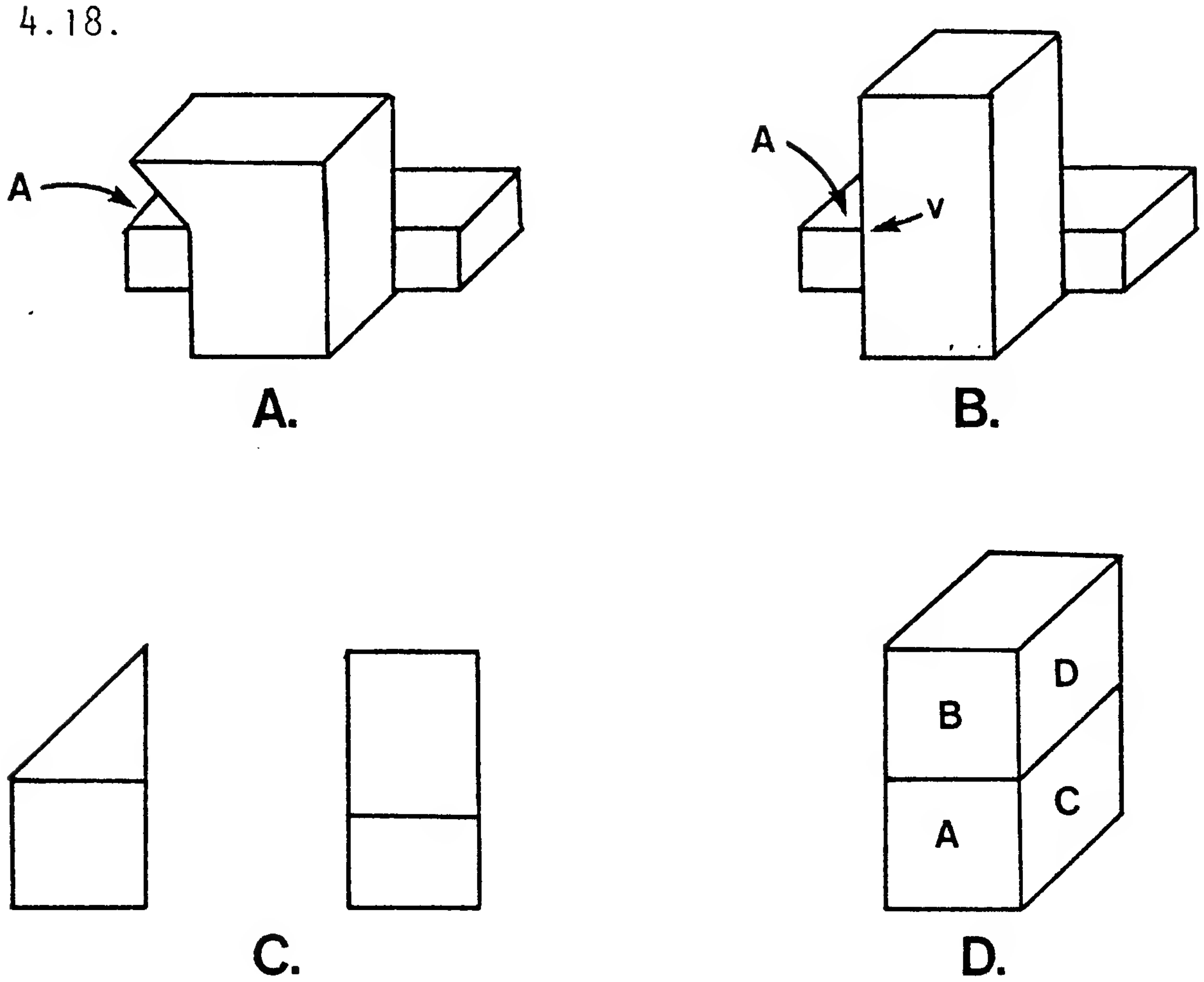


FIGURE 4.19.

parsings, however, one would have to constrain region application: for example, by working in from the background. Thus analysis could start with region A, which borders on the background, rather than with region B, which is completely internal.

OBSTRUCTED MIDPORTIONS AND MISSING LINES REQUIRE SPECIAL TREATMENT

When an object's midportion is obstructed, some form of higher level knowledge is needed to link the endportions. One must be on guard for situations in which endportions can appear unobstructed; for example, region A in figure 4.19A could be recognized as a cross section and projected to form a wedge. This interpretation is not necessarily wrong, just probably less desirable.

Note that the same problem does not arise in figure 4.19B, since the lower right vertex of the corresponding region A can only receive an L1+ label. The reason is that a ray is not allowed to be a collinear extension of the neighboring region lines. With this restriction, the degenerate views of wedge and block in figure 4.19C cannot be recognized. Such views could be recognized by modifying the labeling system to equate T vertices and arrows. However, undesirable interpretations would result, such as finding two objects A-B and C-D in figure 4.19D. It is better to incorporate special knowledge to recognize degenerate views, rather than to change the labeling scheme.

Missing lines disturb scene analysis in various ways. Some missing line situations can be resolved by determining which lines need to be

present to enable a projection. In figure 4.20A region A cannot be projected because the ray between v1 and v2 is missing. Since B cannot be projected either, an impasse is reached. Of the two regions, A seems the better candidate as cross section because missing lines often produce complex regions and A is the simple region. Two rays for a cross section suffice to determine scale change (because scale change is linear), and by projecting a cross section along those rays any missing ones are automatically traced out by the corresponding vertices. Thus projecting A along its two visible rays predicts the line between vertices v1 and v2. Another situations in which missing lines may be detected is at the end of a projection (figure 4.20B).

Postulating missing lines and suggesting new lines to a linefinder are related processes. Insofar as the capability exists to suggest missing lines, this scheme could also help a linefinder.

SOME OBSTRUCTED OBJECTS COULD BE DIRECTLY RECOGNIZED BY PARTIAL PROJECTIONS

One possible inadequacy of the present approach is the inability to directly recognize partially obstructed objects. We are easily able to hypothesize an object with regions A and B in figure 4.20C, regardless of the nature of the obscuring top part. Guzman's[1968] scheme was able to propose a link between A and B because of the arrow, but my "unstacking" procedure cannot act on it until the obscuring part is removed. This inadequacy can be remedied by projecting with incomplete ray data, much as in finding missing lines. Thus region A has one of its rays visible (at

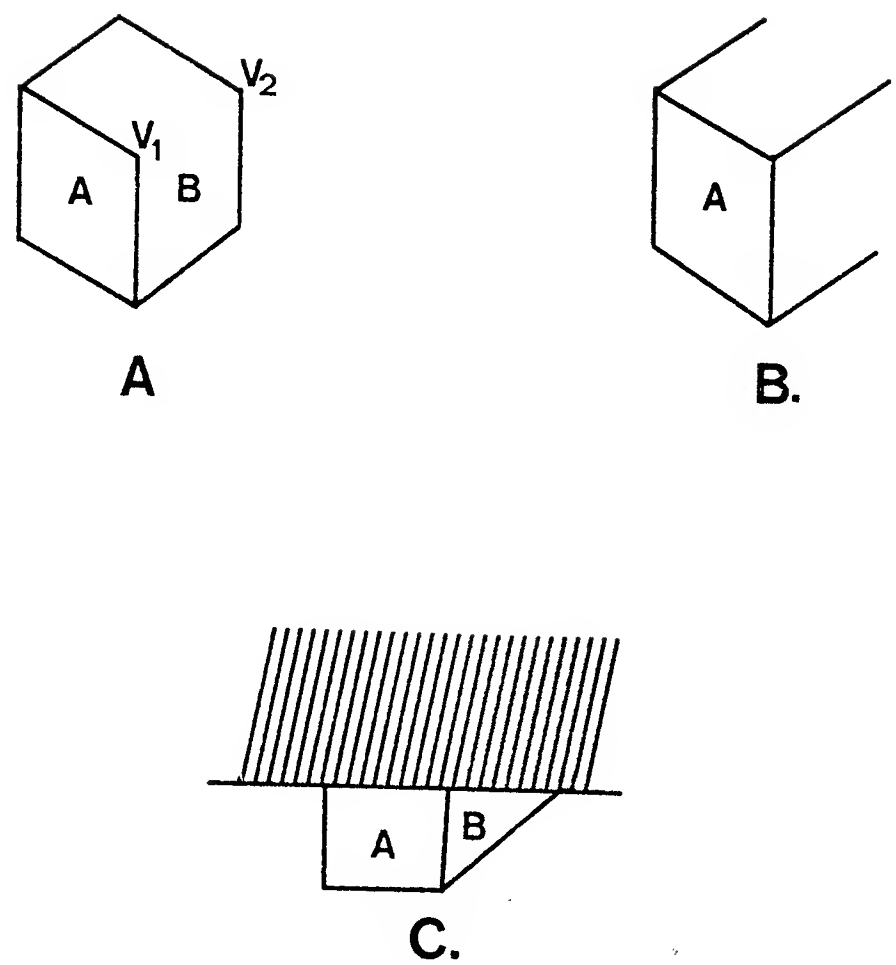


FIGURE 4.20. Missing lines can sometimes be detected.

the arrow vertex), and could conceivably be projected along it if it were suspected the other rays were obscured.

With such a capability, one could go through much of a scene linking regions without removing obstructing objects. Besides aiding missing line conjectures, partial projections would also aid scene reconstruction. For example, partial projection of regions 1 and 3 in figure 4.17E would neatly accomplish reconstruction without using a parallelogram hypothesizer. A reconstructed region may also be required to be something other than a parallelogram, and a partial projection would automatically determine what is needed. However, there are a number of pitfalls that must be avoided in implementing this scheme.

4.4 Other Work on Scene Parsing

This section examines the scene parsing work of Guzman [1968], Rattner [1970], and Waltz [1972]. The present approach provides a standpoint for analyzing what causes their respective techniques to succeed, as well as what causes their failures.

4.4.1 Guzman

Some of Guzman's linking heuristics can be interpreted as an incomplete way of identifying cross sections. The linking of the A1 and A2 regions of an arrow (figure 4.21A) actually hypothesizes that one of the regions, say A1, can be projected along the ray belonging to region A2. The common edge to A1 and A2 sweeps across A2 during the projection, and links A1 and A2 to the same body. The fork linking heuristic, where three links are provided by a fork vertex (figure4.21B), hypothesizes that one of the regions might be projected along the non-region line; in the process, an edge of the projected region is swept on either side of the ray, linking all three regions to the same body.

Guzman augmented his linking scheme with link inhibition, which takes a neighboring vertex into account. Those vertex combinations for which links are inhibited are given in figure 4.21C-G. The T and K inhibitions correspond to the prohibition of T0+ vertices on some FSM transitions.

On the other hand, the L and arrow inhibitions are not generally valid; for example, they prevent correct links for a simple L-shaped object in figure 4.22. The L inhibition represents a banning of an A2+ or F

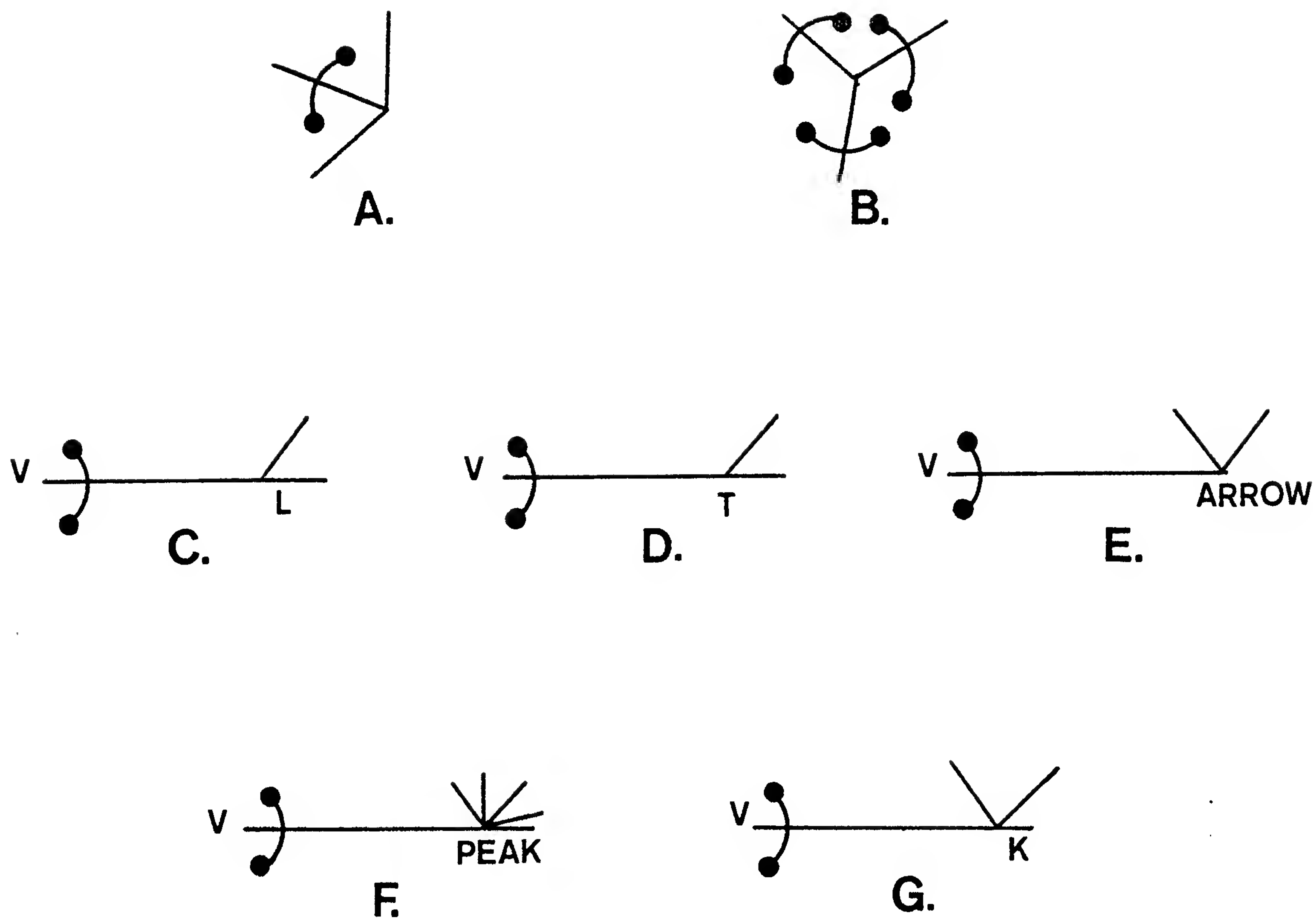


FIGURE 4.21. Guzman links (A - B) and link inhibition (C - G).

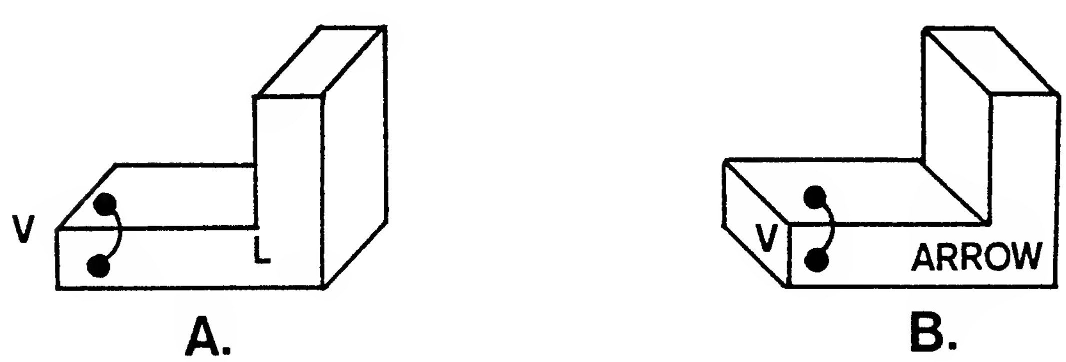


FIGURE 4.22. These link inhibitions are incorrect for these objects.

transition to an L0+ state, and the arrow inhibition represents a banning of an A2+ or F transition to an A0 state. However, all four transitions are allowed in the FSM.

By a conglomeration of weak links, strong links, link inhibition, and region consolidation, Guzman hoped that bad decisions based on a failure to link as well as superfluous links could be averaged out. Instead of this haphazardous accumulation of evidence, the FSM looks at all the vertices around a region to directly provide linkage information. Thus the present approach is an n-vertex approach, where n is the number of vertices around a region, as opposed to a one vertex (link) or two vertex (link inhibition) approach. Since a one or two vertex scheme is not constrained enough it is not surprising that Guzman's approach is too liberal in proposing links.

4.4.2 Rattner

Rattner extends Guzman's approach by the addition of splitting heuristics, which provide an anti-linking scheme in the sense that two regions on either side of a splitting line are hypothesized not to belong to the same body. Rattner provides various heuristics for proposing splits and for extending them through neighboring vertices. His approach is somewhat in the spirit of the present one, in that he decomposes a vertex into two adjacent regions that might belong to the same body and into other unlinked regions. This is like selecting a region and ray, and assigning everything else to the "+" category.

He first designates some vertices as splitting vertices (figure 4.23A), from which a split is initially obtained or through which a split is propagated. The general 4-line vertex, for example, can receive either of the splits in figure 4.23B; the choice of split is decided by context.

By splitting three adjacent lines from the rest, Rattner is actually decomposing vertices. It is this decomposition which allows his approach to handle alignment as well as it does. His approach is not general, though, because he can handle at most a 5-line vertex. In this sense the "+" assignment in the FSM forms the most general splitting heuristic, as it represents arbitrarily many lines. His heuristics, moreover, are fairly ad hoc and too local, and hence do not apply in many situations. As in Guzman's approach, he is counting on a vague global compilation of evidence to result in a unique parsing.

Many of his heuristics can be derived as special cases of the FSM

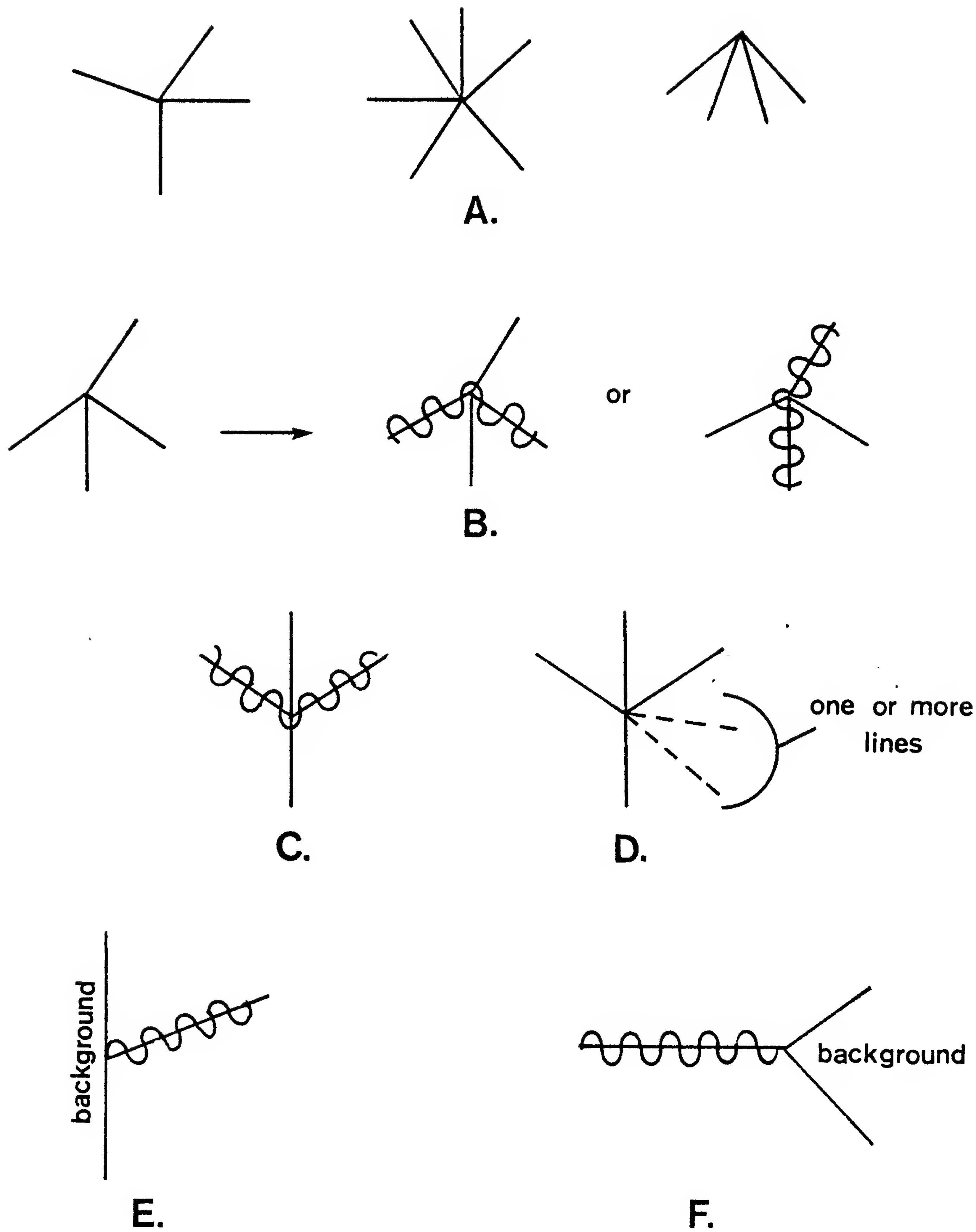


FIGURE 4.23.

operation. Here are some examples.

1. Internal T (figure 4.23E)

When a T is formed with a T0 background region, the shaft receives a split. The T0 background means that the two collinear parts are separate edges, rather than a single edge of some obstruction. The FSM would assign L1+ to the T1 and T2 vertices, so that a hypothetical projection of either the T1 or T2 region would move away from, rather than encompass, the other region. Hence the split is justified.

2. Special multi (figure 4.23C)

This heuristic is like the last one, with the background replaced by a region divided by a line forming an arrow with the shaft of the T. The only possible assignment to the upper left and upper right portions of the vertex (other that trivial L1+ ones) are respectively A1+ and A2+. Either links the two upper regions, and hence suggests a split along the non-collinear edges.

He provides a more general form of this heuristic (figure 4.23D) which comes close to the "+" assignment. My impression though is that he does not make use of this flexibility, since his initial vertex specification allows for at most 5 lines.

3. Split to external concavities (figure 4.23F)

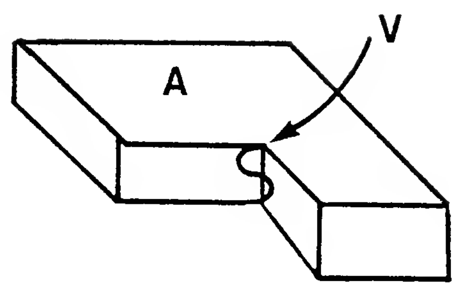
Both lines of the fork bordering on the background are true edges of the corresponding regions. Projecting one region along the opposite edge would fill in part of the background; thus neither region can act as a cross section. They may still belong to the same body, however, as

evidenced in figure 4.24A where they are encompassed by projection of region A. Rattner eliminates this case by requiring v to be a splitting vertex. Even this restriction does not render the heuristic foolproof, because v could result from an accidental alignment as in figure 4.24B.

4. Split between pairs of splitting vertices

This is another heuristic that fails to apply generally. There are a number of ways to combine splitting vertices so that the two regions on either side of their common line belong to the same body. An A1+ or A2+ assignment can be found for any of the splitting vertices, and the transitions A1+ -> A2+, A2+ -> A1+ are allowed in the FSM.

To summarize, Rattner's approach deals well with alignment because of his system's ability to decompose vertices. His heuristics, however, are a haphazardous collection of local observations: sometimes they work, sometimes not. He finds alternate interpretations for an ambiguous scene by throwing out one heuristic and trying an alternate one. This does not always work because his heuristics are incomplete. Thus he is able to find the "normal" interpretation of figure 4.18, but not the alternate one.



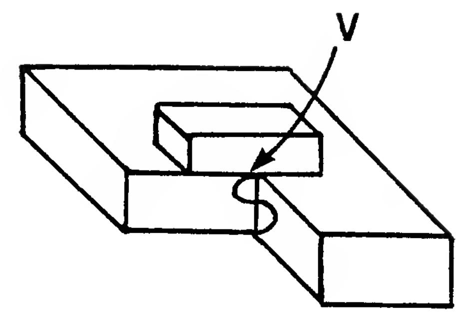


FIGURE 4.24. Incorrect splits generated by Rattner's program.

4.4.3 Waltz

The labeling approach of Waltz is based on an exhaustive enumeration of vertex types in terms of various edge labels, such as convex, concave, obscuring, crack, and shadow. By adding these labels to the basic Huffman set, Waltz was able to expand Huffman's subdomain of allowed polyhedral scenes while maintaining a favorable ratio between realizable versus possible junctions. Huffman's subdomain was restricted to scenes of trihedral polyhedra in general position and without alignment of any form. Waltz's extension covered shadows and trihedral alignment, which is a form of stacked alignment where only three distinct planes meet at a junction; for example, junction x in figure 4.25 shows trihedral alignment, while junction v represents non-trihedral alignment.

In this expanded but still restricted domain, the number of topological junction types is rather small. No junction of more than six lines may appear, which happens when all edges of a three plane junction are visible. Within a particular junction type, the proportion of realizable junctions is also rather small. The incorporation of shadow lines actually decreased the proportion of realizable junctions for a particular type. One thrust of Waltz's work was showing how region illumination and orientation impose severe restrictions on junctions with shadows.

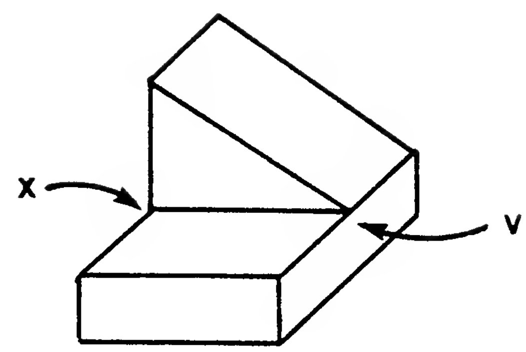


FIGURE 4.25.

EXPANDING WALTZ'S SUBDOMAIN EXPLODES REALIZABLE JUNCTION TYPES

Successful in this particular subdomain, Waltz sought to generalize his labeling scheme to handle non-trihedral and accidental alignment. He introduced several new line labels and gave what he considered to be the most common junctions produced by them. Unfortunately, this begins to take on an ad hoc flavor, and it is not hard to concoct simple examples which contain junctions for which he has no labeled type (e.g., v in figure 4.25). He is evidently wary of including such alignment junctions in his regular data base because they might interfere with labeling of scenes without such alignment.

There is no evidence that the labeling scheme generalizes outside Waltz's subdomain. The essential problem is that arbitrary alignment greatly explodes the number of junction possibilities. First, junctions with arbitrarily many lines may appear. It is very easy to create examples with 7,8,9 or more lines. Second, the number of realizable junctions within a particular type increases enormously. Whereas there are only 16 distinct K junctions in his original subdomain, Waltz notes that accidental alignment of one object edge with a vertex results in 18,000 new realizable K junctions. This number is only for one particular form of alignment, and there are other ways new junctions can be created: junction to junction accidental alignment, accidental alignment with a junction already formed by accidental alignment, etc. Indeed, the problem seems more severe, the more lines a junction has.

The explosion results from a multiplicative effect. The two junctions

or junction and line coming together in an accidental alignment are independent of one another, i.e., one does not constrain the other much. Thus the number of realizable junctions in a junction-junction alignment is roughly the product of the number of ways one junction might appear with the number of ways the second junction can appear. If the alignment is an obscuring edge falling on a junction, then the number of realizable junctions is the number of ways this edge can combine with the junction. The large number of realizable possibilities makes it for all practical purposes impossible either to enumerate them or to label a scene with the augmented set.

Waltz puts forth several arguments minimizing the importance of handling arbitrary alignment. He notes that accidental alignment can be resolved simply by moving with respect to the scene. He argues further that many types of alignment are extremely unlikely, and hence there is no great need to make provisions for them.

Counterarguments can be given against these points. To be sure, accidental alignment can be resolved by moving. But when people look at two-dimensional line drawings, movement is not helpful. I do not think people have any great difficulty handling alignment, whether they are confronted with a drawing or with an "unlikely" form of alignment. In particular, I don't think people apply a totally different mechanism towards aligned scenes than they do to more restricted scene types. The approach I have presented, on the other hand, works without modification on arbitrary alignment. No special provision is made to handle new junction

types; 10-line junctions are dealt with as readily as 2-line junctions.

I believe the failure of the labeling approach with alignment lies with the need to label all lines of a junction. The success of the present approach and of Rattner's with alignment is due to a decomposition of complex junctions, so that one need account for only a few lines of such junctions. The FSM works on a very local basis, picking out single objects while ignoring the environs. Once a few lines have been interpreted as part of an object, the junction becomes less ambiguous and complex.

Waltz expressed satisfaction that his scheme works without the need for locating hidden lines or regions, and in his subdomain he is correct. But to handle arbitrary alignment, such an estimation could well be needed if the present approach is any guide. For, in the reconstruction phase hypotheses are made about hidden lines and regions, while in cross section selection hidden rays are hypothesized at L0+ and L1+ vertices.

4.5 Single Cylinder Description

The axis of a polyhedral cylinder is straight, and hence merely needs a symbolic length description. A polyhedral cylinder is most like a cone or cylinder prototype, and their height-width quantizations can serve this purpose. The width here might correspond to the maximum width of the cross section.

The scale change function is linear, and also receives a qualitative description: "stays the same" (zero scale change), "grows or shrinks slowly", or "grows or shrinks rapidly". Taking a cue from the cone/cylinder distinction, the boundary between "stays the same" and "grows or shrinks slowly" is set at an angle of 30 degrees at the point where the rays would meet when extended. A 90 degree boundary between "grows or shrinks slowly" and "grows or shrinks rapidly" is consistent with the distinction between a high and low angle slant.

CROSS SECTIONS ARE DESCRIBED IN TERMS OF PROTOTYPES AND MODIFIERS

As mentioned earlier, simple geometric shapes make good prototypes. Block and wedge, for example, can be expressed as projections of rectangles and triangles. Indentations and protrusions serve as the two types of modifiers for all prototypes.

For regular or systematic modifications, a group modifier such as jagged or saw-toothed is more appropriate than individual modification description, but a study of such modifiers has not been carried out here. Neither has the problem of cross section segmentation been addressed;

sometimes a cross section needs to be segmented and described by two or more prototypes, such as the hexagon-square combination in figure 4.26.

AN EXPERIMENT IN MODIFICATION ASSESSMENT

I have studied the problem of prototype selection and modification for the case of square versus rectangle. A rectangle or square under an arbitrary projection into 2-space rarely appears as a rectangle of course, but as a parallelogram when there is no perspective deformation, and as a trapezoid or trapezidium with deformation. Auxiliary considerations are required to equate deformed regions with prototypes, but the present study presumes no projective deformation has taken place.

Systematic modifications were made to a square to yield a variety of objects. Members of the AI Laboratory were asked to categorize each object either as: a rectangle or square modified by an indentation (I), a rectangle or square modified by protrusions (P), or an object not well described by these alternatives (N for neither). Sample results on some objects are presented in figure 4.27.

A simple parameterization was devised that sorted these objects correctly into the above three groups. A plot of

$$\frac{\text{protrusion depth}}{\text{square height}} \quad \text{vs.} \quad \frac{\text{indentation gap}}{\text{protrusion breadth}}$$

is presented in figure 4.28. When two or more protrusions emanated from the side of a rectangle, the parameters were obtained by considering only the largest.

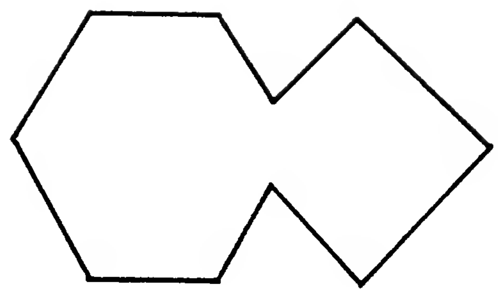


FIGURE 4.26. This cross section is best described by two prototypes: hexagon and square.

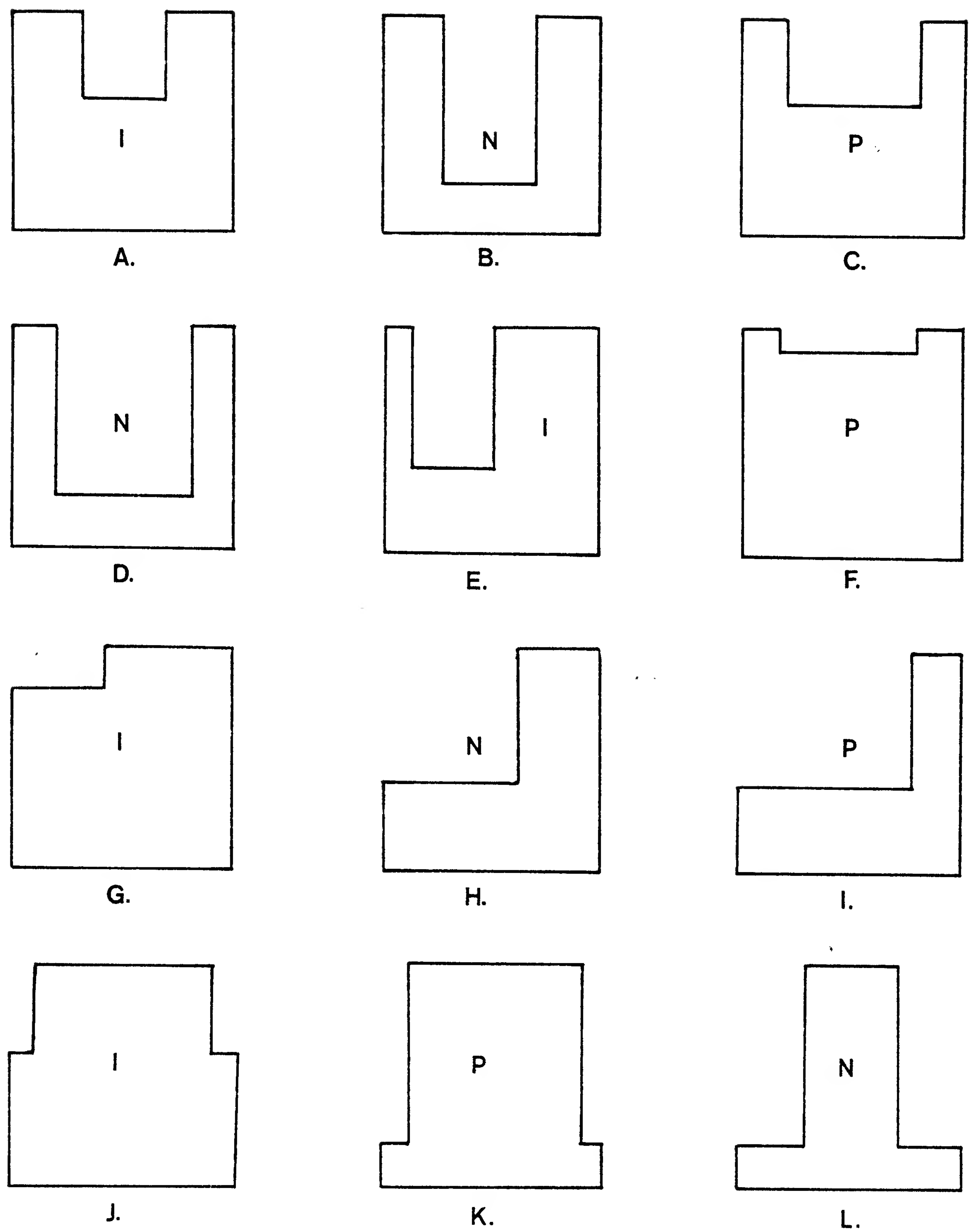


FIGURE 4.27.

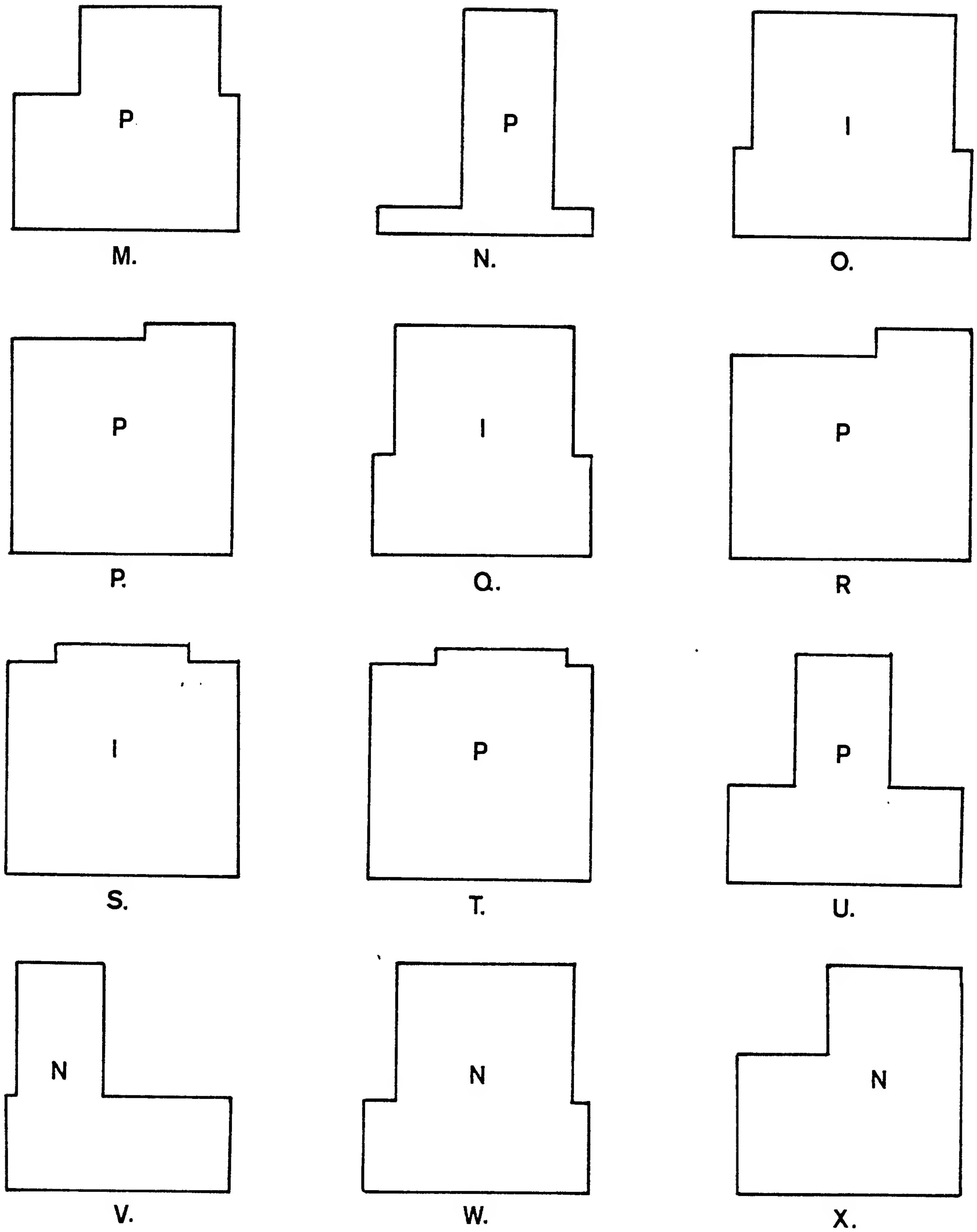


FIGURE 4.27. Continued.

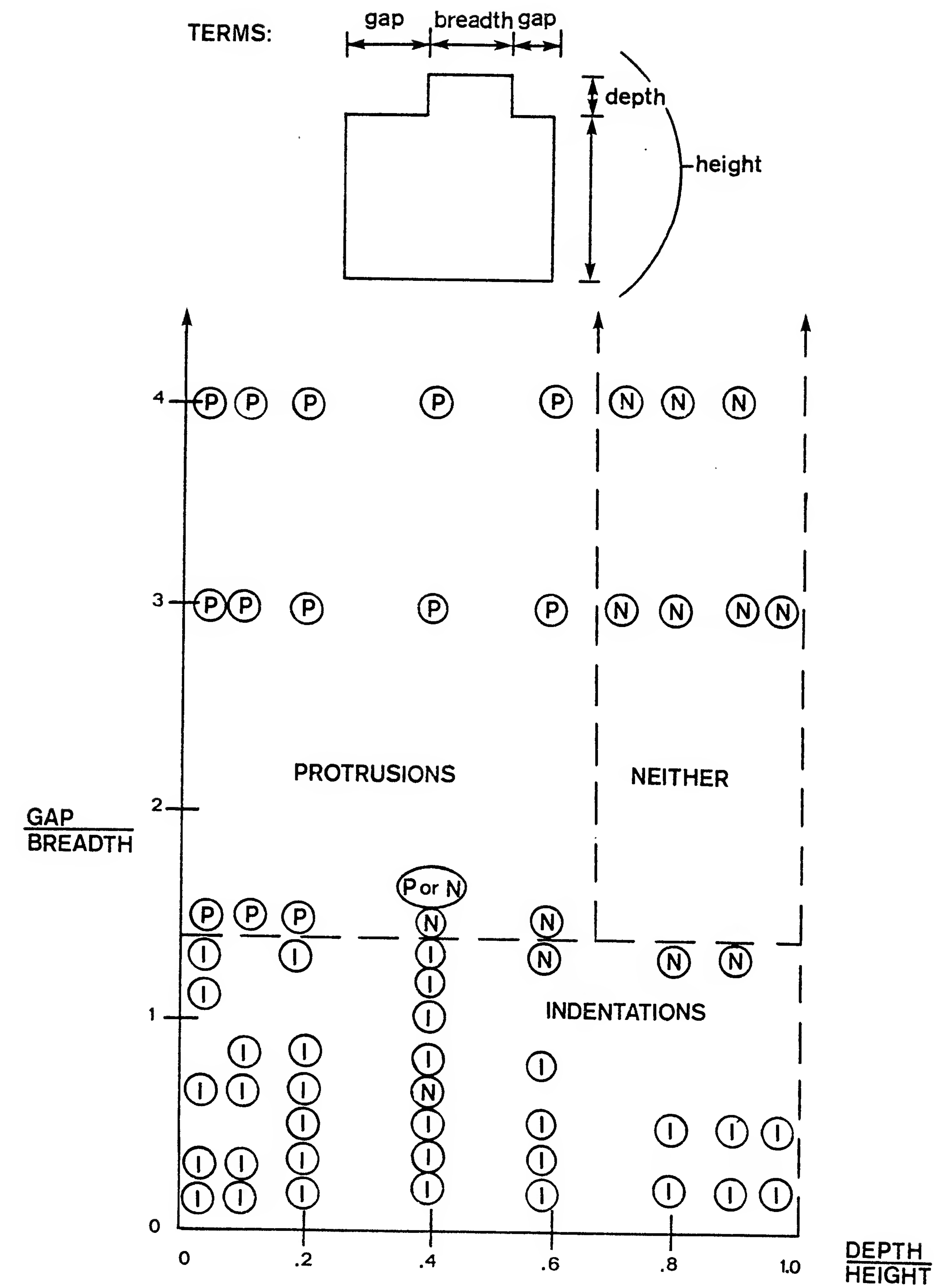

TERMS:
gap
breadth
gap
depth
height
PROTRUSIONS
NEITHER
INDENTATIONS
P or N
GAP / BREADTH
DEPTH / HEIGHT
0
.2
.4
.6
.8
1.0
1
2
3
4

FIGURE 4.28.

This parameterization has a simple interpretation. Protrusions must be sufficiently isolated from the rest of the object to resist integration as part of an indentation, which happens when the gap:breadth ratio becomes large enough. Yet the protrusion must not be so large as to become significant in size to the rest of the object, as when the depth:height ratio approaches one. When the latter ratio is near one, the object is composed of at least two roughly equal and distinct pieces, and hence receives an N categorization. An N categorization implies either the need for a more complex prototype or a need for segmentation into two or more prototypes. Thus objects B and D might be said to be U-shaped while object L is an inverted T, whereas objects V and W might be best described as one rectangle atop another.

SOME ANOMALIES ARE EXPLAINED BY SYMMETRY

There are some anomalies to this parameterization, but they disappear when the simplifying effect of symmetry is taken into account. For example, the protrusions of objects O and P, Q and R, and S and T, respectively, are proportionately equal. Yet the symmetry in objects O, Q and S causes the gaps to be seen as indentations, while the asymmetry of P, R and T causes a protrusion interpretation.

Another discrepancy is between objects U and V, and between Q and W. Once again, the top protrusions are of proportionately equal size, yet in one case the protrusion is symmetrically placed and in the other it is not. They were interpreted, respectively, as P and N. A final mystifying result

was obtained for X, which because of its symmetrical shape resisted decomposition.

The conclusion to be drawn from these anomalies is that in describing a feature, symmetry favors I over P and P over N. The preference of P over N means that we are more likely to interpret a feature as a modification than to segment the object into two or more separate but equal pieces.

Irregularities of outline also lead to I over P preference. Thus the top protuberances in figure 4.29 are approximately the same size as in figure 4.27C, but the modification looks like an indentation rather than two protrusions. Protrusions can be considered as constructive additions, indentations as destructive subtractions. A constructive addition leads to more regular objects than a destructive subtraction, which tends to leave irregular pieces. Indeed, the square in figure 4.29 looks as though a gouge had been made in the top.

THE LYING VERSUS STANDING BRICK PROBLEM IS REEXAMINED

A special problem in cross section selection is presented by the rectangular block. Any of its faces could serve as cross section, but sometimes one choice seems more appropriate than another. A square face when it exists intuitively seems the best choice as cross section (figure 4.30A). When there are no squares, let us assume that the face most like a square, namely one whose length ratio of shorter to longer side is closest to 1, is the appropriate choice (figure 4.30B and C). This simple assumption also corresponds to intuition, and can be used to derive in an

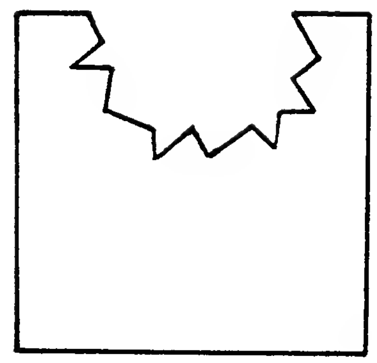


FIGURE 4.29. These jagged edges cause an indentation interpretation.

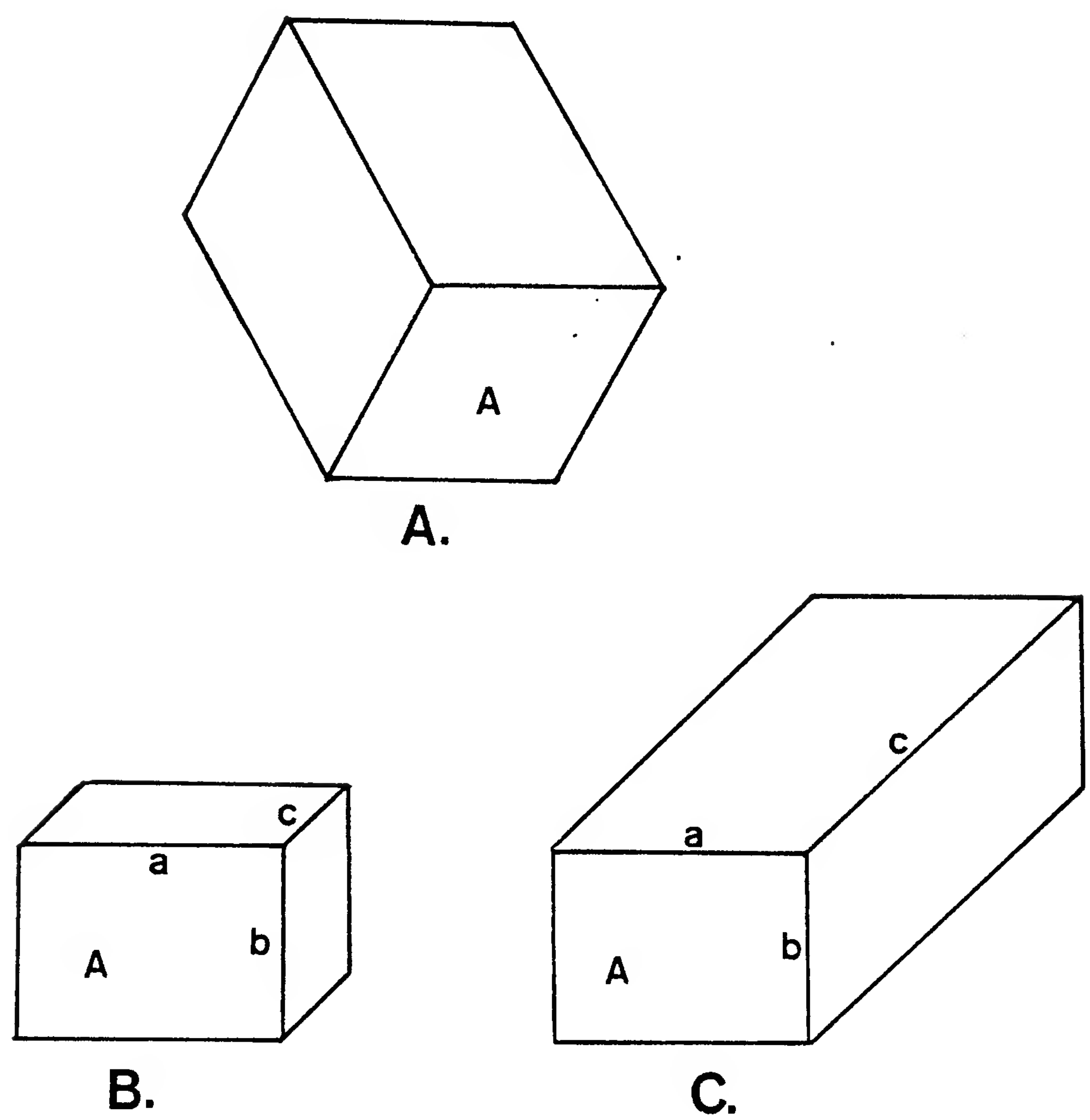


FIGURE 4.30. The most square-like region is chosen as cross section for block.

alternate manner some results on standing and lying bricks obtained by Finin [1971].

Suppose region A in figures 4.30B-C is the most square-like. Its shorter side is of length b, its longer side is of length a, and its ray is of length c. Let us consider the restrictions imposed by the choice of A on the range of possible values of c. At the lower range (figure 4.30B),

$$c/b < b/a \quad \text{or} \quad 1 < b^2/ac$$

Otherwise the region with sides b and c would be more square-like. At the upper range (figure 4.30C),

$$c/a > a/b \quad \text{or} \quad 1 > a^2/bc$$

This is precisely Finin's parameter y^2/xz for distinguishing a lying from a standing brick.

When $1 < b^2/ac$ let us say the brick is short; when $1 > a^2/bc$ we say the brick is long. Suppose the cross section is a vertical face of a brick. If the brick is short, it is standing; if long, it is lying. Suppose the cross section is a horizontal face of a brick. If the brick is short, it is lying; if long, it is standing. This exactly duplicates Finin's results. To summarize:

	short	long
vertical	standing	lying
horizontal	lying	standing

4.6 Complex Objects

A complex object is one that cannot be described as an unmodified cylinder. Such objects are segmented into single cylinders, each of which is modified with indentations and protrusions as required. Any complex object can ordinarily be segmented in a variety of ways, each way yielding a different description, and the problem is to select the best one.

Segmentation is accomplished by projecting some cross section to form a single cylinder. During the projection, protrusions and other cylinders are segmented while indentations are filled in. These features are signaled by obstructions or barriers to the path of the projection. Different descriptions result from using alternate cross sections and by interpreting a modification differently (i.e., the distinction between indentations and protrusions is sometimes equivocal).

Cylinder modifications are also solid objects, and can themselves be described in cylinder terms. Protrusions and indentations for cylinders are related to their two-dimensional counterparts by projection: a modification to a region will yield the same type of modification in three dimensions when the region is projected. Since modifications are solid, they are distinguished from separate cylinders only by size. Hence they are treated equally during segmentation, which can now be conducted on the simplified assumption that a complex object consists of one main cylinder with modifications. Modifications can be sorted by size from separate cylinders in later stages of processing.

FIRST, THE MAIN CYLINDER AND ITS CROSS SECTION ARE DETERMINED

A simplified procedure that carries out the complete description task is outlined in figure 4.31. To begin, the main cylinder of an object is obtained by finding the region with greatest area; in figure 4.32A this corresponds to region A. This region is hypothesized to be a part of the main cylinder because the main cylinder is usually the largest component of an object and because the largest region is a good indicator of the largest component. Unfortunately, projective deformation may result in a region with apparent largest area being actually smaller than some foreshortened region, yet as a first approximation apparent largest area is a reasonable choice that evidently corresponds to human size judgment.

A cross section for the main cylinder is chosen next, and is restricted to be either the largest region A or its most complex bordering region B. This restriction rests on the observation that the most complex region often gives the best characterization of an object; it has the additional effect of limiting the segmentation possibilities to a manageable number.

With regard to definitions, one region borders another if they share a convex edge in the Huffman sense. Not all regions sharing an edge with A need actually border A in 3-space, since the shared edge may be obscuring. Complexity is defined on the basis of number of sides and of region regularity as follows:

1. If A is a triangle and B is a quadrilateral, then A is more complex than B.

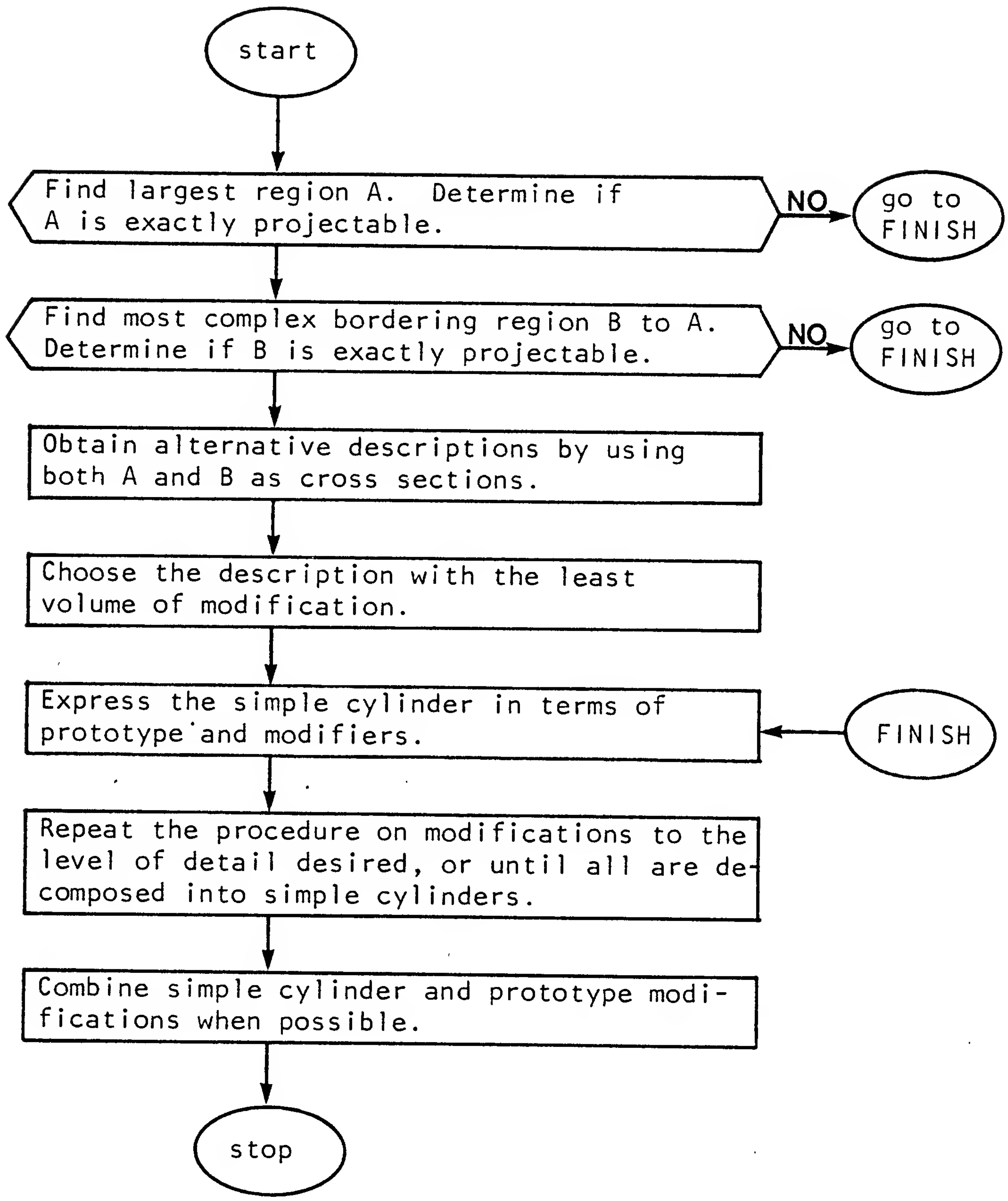

start
Find largest region A. Determine if A is exactly projectable.
NO
go to FINISH
Find most complex bordering region B to A. Determine if B is exactly projectable.
NO
go to FINISH
Obtain alternative descriptions by using both A and B as cross sections.
Choose the description with the least volume of modification.
Express the simple cylinder in terms of prototype and modifiers.
FINISH
Repeat the procedure on modifications to the level of detail desired, or until all are decomposed into simple cylinders.
Combine simple cylinder and prototype modifications when possible.
stop

FIGURE 4.31.

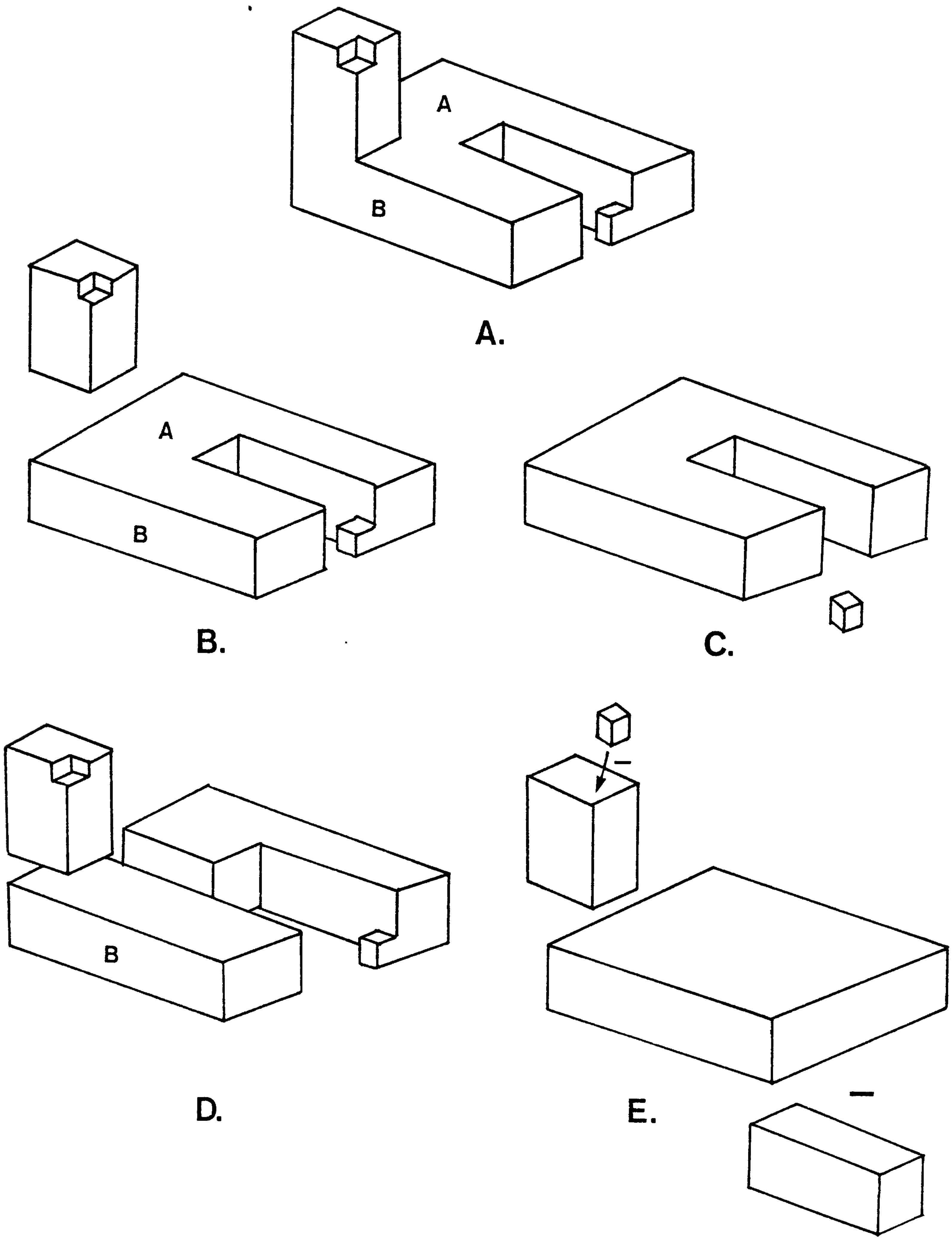


FIGURE 4.32.

2. Otherwise, if A has more sides than B,
   then A is more complex than B.

3. When A and B have the same number of sides,
   the more regular region is less complex.

Regularity will not be precisely defined here, although the ordering it induces on some region types is fairly clear. For example, the ordering of 4-sided regions would probably be square, rectangle, rhombus, paralellogram, trapezoid, and trapezidium. Triangles are judged more complex than quadrilaterals because they are preferential as cross section in wedge shaped objects.

Neither A nor B would yield a single unmodified cylinder in a projection, and to choose between them it is necessary to compare the amount of modification in the respective descriptions that they generate.

In obtaining the first description with region A, a protrusion must be removed to render A projectable (figure 4.32B). As A is now projected, it encounters a barrier in the lower right portion. A decision must be made at this point to terminate or to continue the projection. The latter choice is more appropriate here, and leads to the segmentation of a small protrusion as in figure 4.32C. Using B as cross section, the decomposition in figure 4.32D is eventually obtained (discussed in the next section).

Comparing the two descriptions, clearly A is the better cross section because its generated description requires less modification than B's. A way of measuring the amount of modification to a description is by summing the volume from each indentation and protrusion. A rough volume estimate for a modification could be obtained by multiplying the area of the largest

region against the average length of its rays.

THE INITIAL DESCRIPTION IS REFINED AND REWORKED

After this initial phase, the main cylinder is described in terms of prototype and modifiers by examining cross section shape. For example, the main cylinder in figure 4.32C is described as a rectangular block with side indentation (figure 4.32E). The individual modifications are similarly described by running them through the same procedure. Thus the protrusion removed in figure 4.32B is described as a block with indentation in the corner (figure 4.32E). Modifications can themselves form complex objects, and one could conceivably run them and their own modifications recursively through the procedure until everything is decomposed into single cylinders. Or one could stop this process at some coarser level of description, using the current main cylinder description and disregarding its modifications.

In the last step of the procedure, an attempt is made to simplify the description by combining some subparts into one part. Due to inadequacies in the procedure or to the vagaries of modification, an object may be segmented into too many pieces. For example, a notch in the L-shaped block in figure 4.33 has dissected the L region. As a result the procedure comes up with the indicated description because it chose A as cross section. The last step recombines the rectangular block and protrusion into the preferable L-shaped prototype.

The next section goes into greater detail with various aspects of this routine.

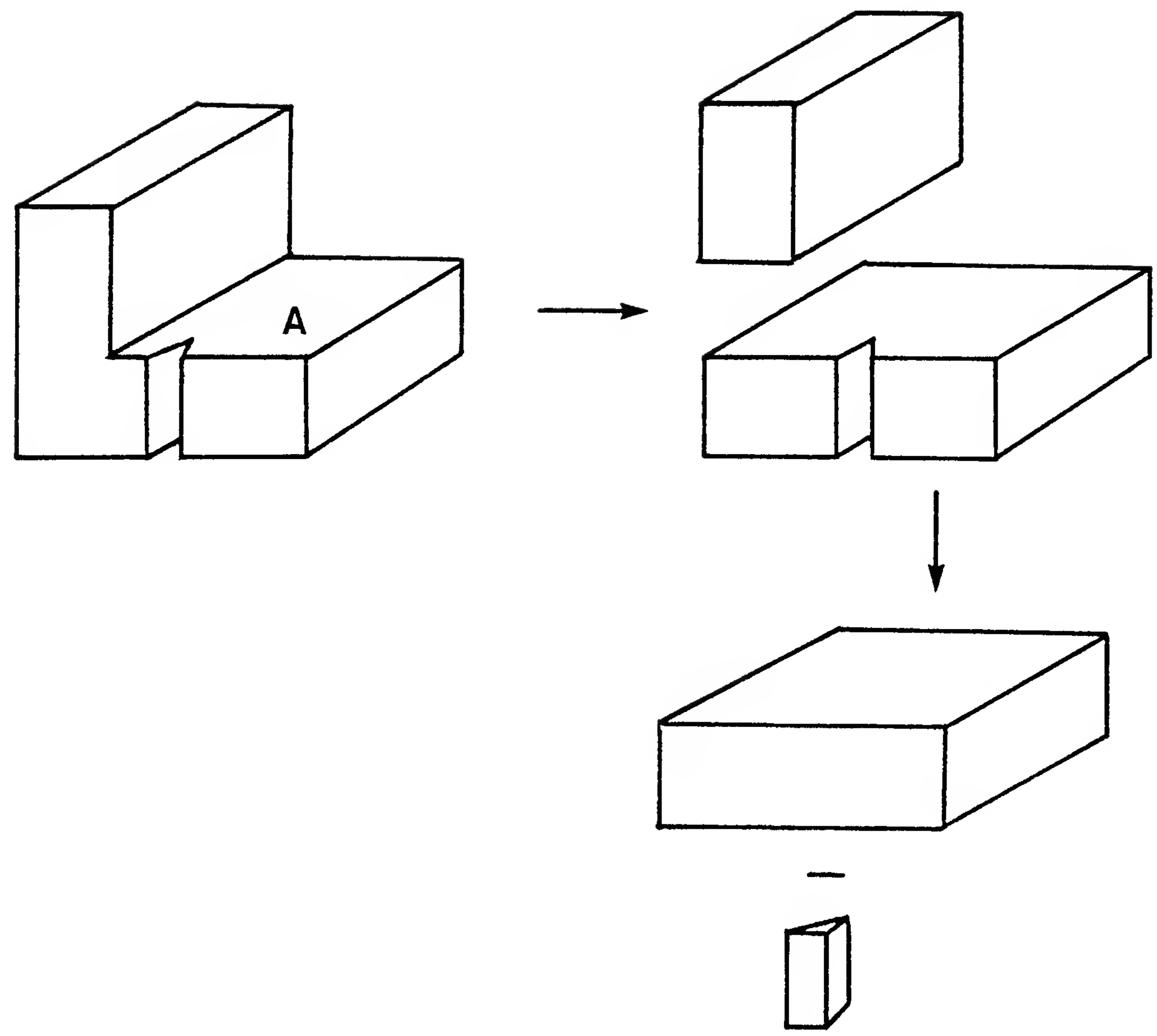


FIGURE 4.33.

4.6.1 A More Detailed Examination

The real difficulty in applying the procedure lies in the third step, where a region is projected to yield main cylinder with modifications. The present section outlines one way this could be accomplished.

A region must first satisfy the conditions for cross sections before it can be projected. This may involve segmenting a protrusion so as to remove a concave edge or an obscuring edge that forms a forbidden T0+ vertex. Segmentation is accomplished by locating a junction composed of a convex edge of the region and one of these concave or obscuring edges, and by extending this convex edge through the junction and across the protrusion region. When applied to region A of figure 4.32A, the extra line in figure 4.34A results. The segmentation routine then removes the protrusion, which now looks like a separate object, and reconstructs region A.

When a protrusion has two or more regions which need a segmentation line, as in 4.34B, a partial projection could supply the remaining lines after the convex edge extension. Thus newly formed region C, a result of the first line extension from A, could be projected to add the other line as the trace of the indicated vertex. Since C has two visible rays, its scale change function is completely determined, and the missing ray will be properly placed.

Once all protrusions have been removed, a region may qualify as cross section but not have a strictly linear scale change. Thus cross section A in figure 4.34C has zero scale change for every edge except e1 and e2;

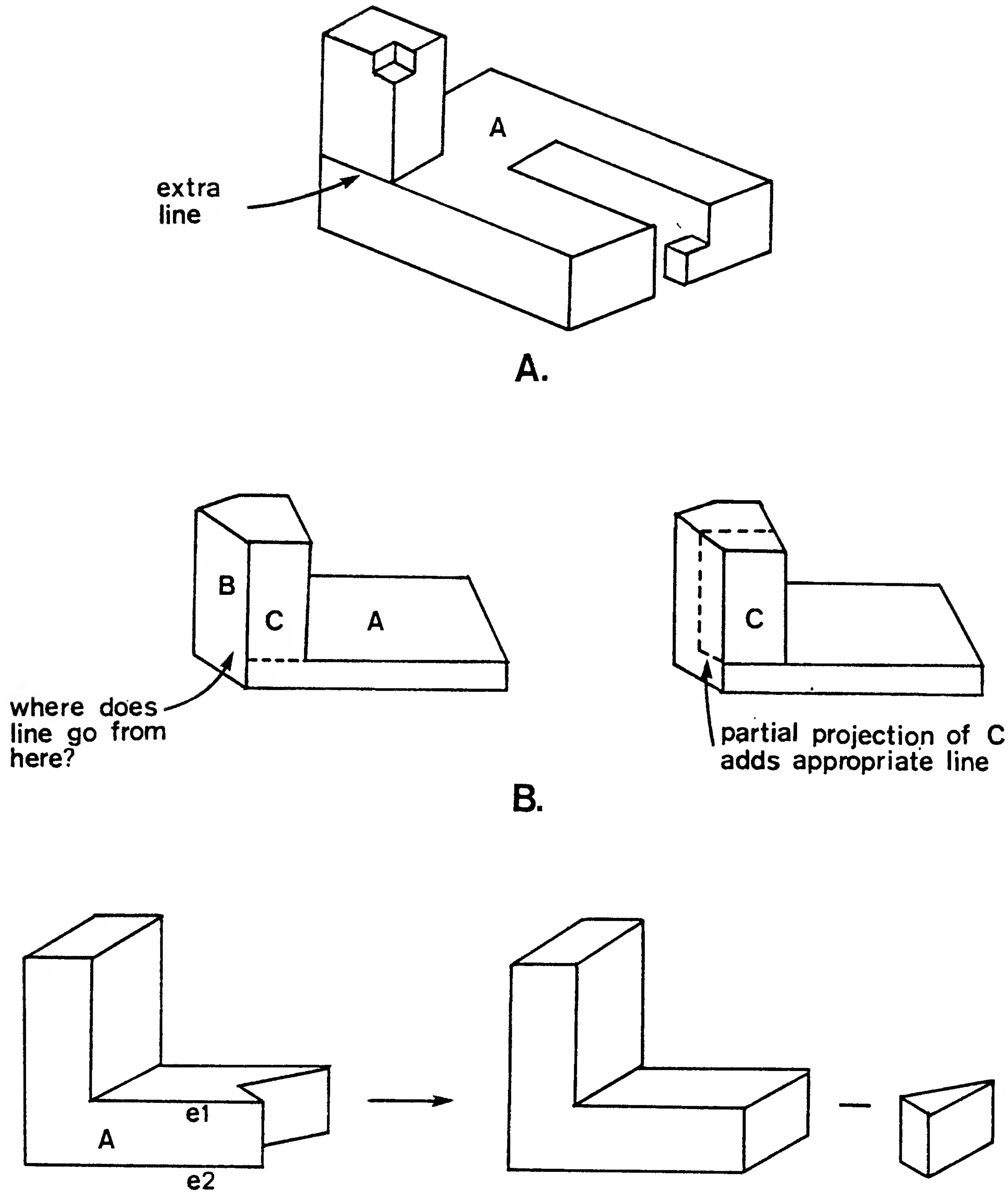

A
extra
line
A.
B
C
A
where does line go from here?
C
partial projection of C adds appropriate line
B.
e1
A
e2
C.

FIGURE 4.34.

hence the object is not a simple cylinder. To form the main cylinder, one scale change value should be chosen for all the edges, zero being the best value in this instance. If A is now projected, the result is an L-shaped block with indentation, since the indentation is "filled in" during projection. Modifications in general are indicated by non-uniform scale change because of the perturbations they cause on rays.

BARRIERS TO A PROJECTION MAY TERMINATE IT OR MAY BE BYPASSED

Once a projection begins, it may run into a barrier that prevents exact projection. Two decisions then have to be made: (1) to stop or continue the projection, and (2) to classify the barrier as an indentation or protrusion. To facilitate in these decisions, barriers are subdivided into interior and exterior barriers. Interior barriers are interruptions within the borders of the projecting cross section, as for the objects with cross section A in figures 4.35-7, while exterior barriers lie outside the border, as for objects in figure 4.38. Exterior barriers are distinguished from interior ones by the presence of a concave region angle at a shortened ray. These subdivisions are considered separately below.

Internal Barriers. These barriers represent a removal of material from the main cylinder being formed by projection, a removal that results in a decrease of cross section area. When too much material is removed, the cylinder loses its integrity and hence should be segmented at that point. This yields a protrusion interpretation for the barrier. Otherwise

the cross section continues its projection, perhaps in a modified form as discussed later. A 50% decrease in cross section area serves as dividing line between these two possibilities, and is consistent with the distinction between indentations and protrusions in figure 4.28.

At the point of the barrier, it is therefore necessary to gauge how much of the cross section is being decreased. This can be done by completing the barrier edge portions on the cross section with projections. For example, figure 4.35B shows cross section A at the point where it reaches the barrier. Region B is now projected as in figure 4.35C to divide A into barrier and remaining cross section portions, as in figure 4.35D. Clearly the barrier portion comprises more than 50% of the area of A. Hence the projection stops at that point, leading to the protrusion segmentation in figure 4.35E.

When the barrier proportion is less than 50%, projection does not cease at that point. In figure 4.36 projection of A continues past the barrier to yield an indentation. Note that barrier completion is a little more complex here, since the barrier shares two edge portions of A. Hence both regions B and C are projected to complete the barrier, as in figure 4.36B where only the projection lines that lie on A have been depicted. The barrier portion is formed from the intersection of the projection lines, as in figure 4.36C.

Subjectively speaking, the barrier portion in figure 4.36C seems to lie inside the region, as if it were a missing chunk; hence it is interpreted as indentation. On the other hand, the barrier portion in

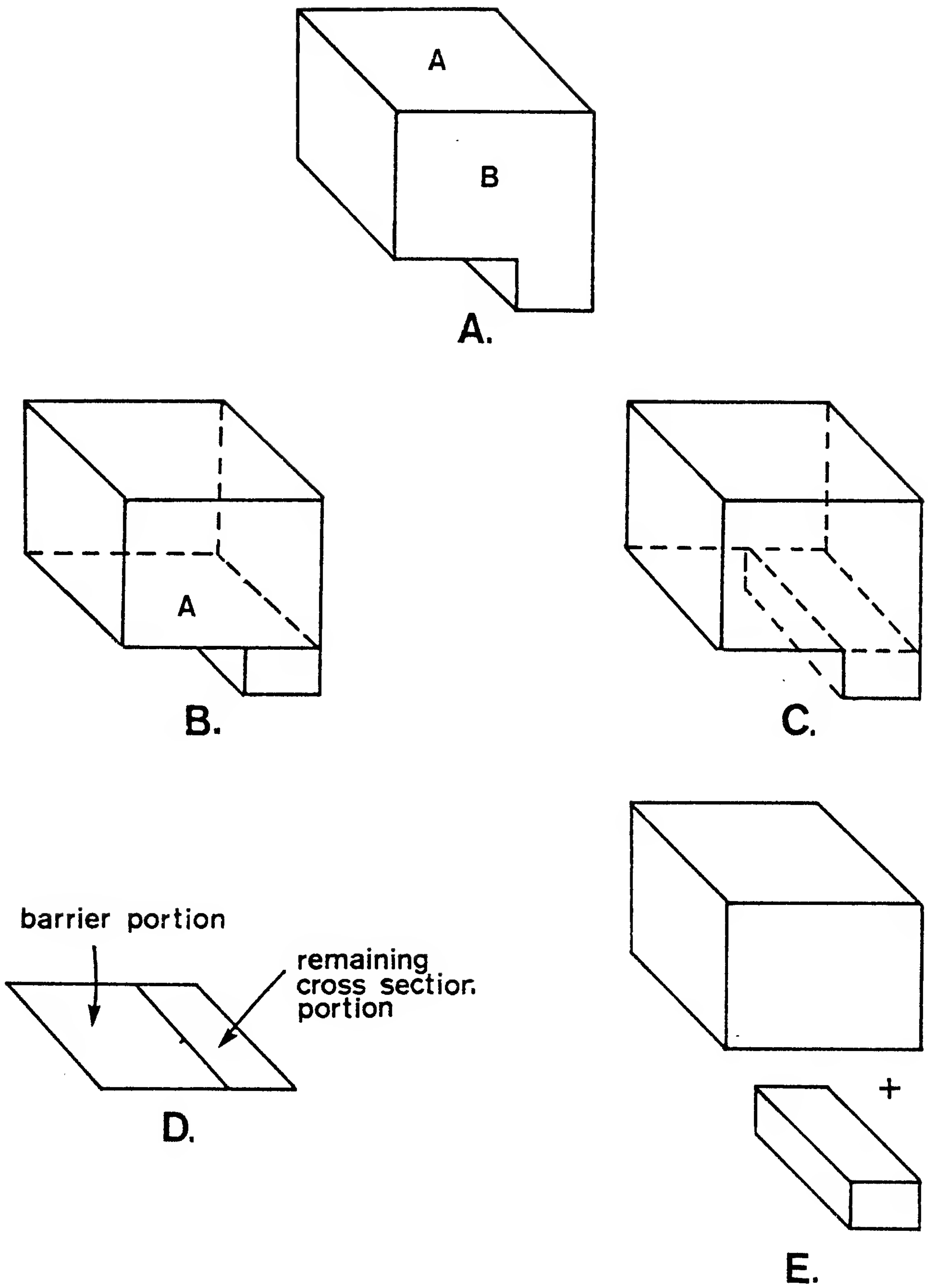

A
B
A.
A
B.
C.
barrier portion
remaining
cross section
portion
D.
+
E.

FIGURE 4.35.

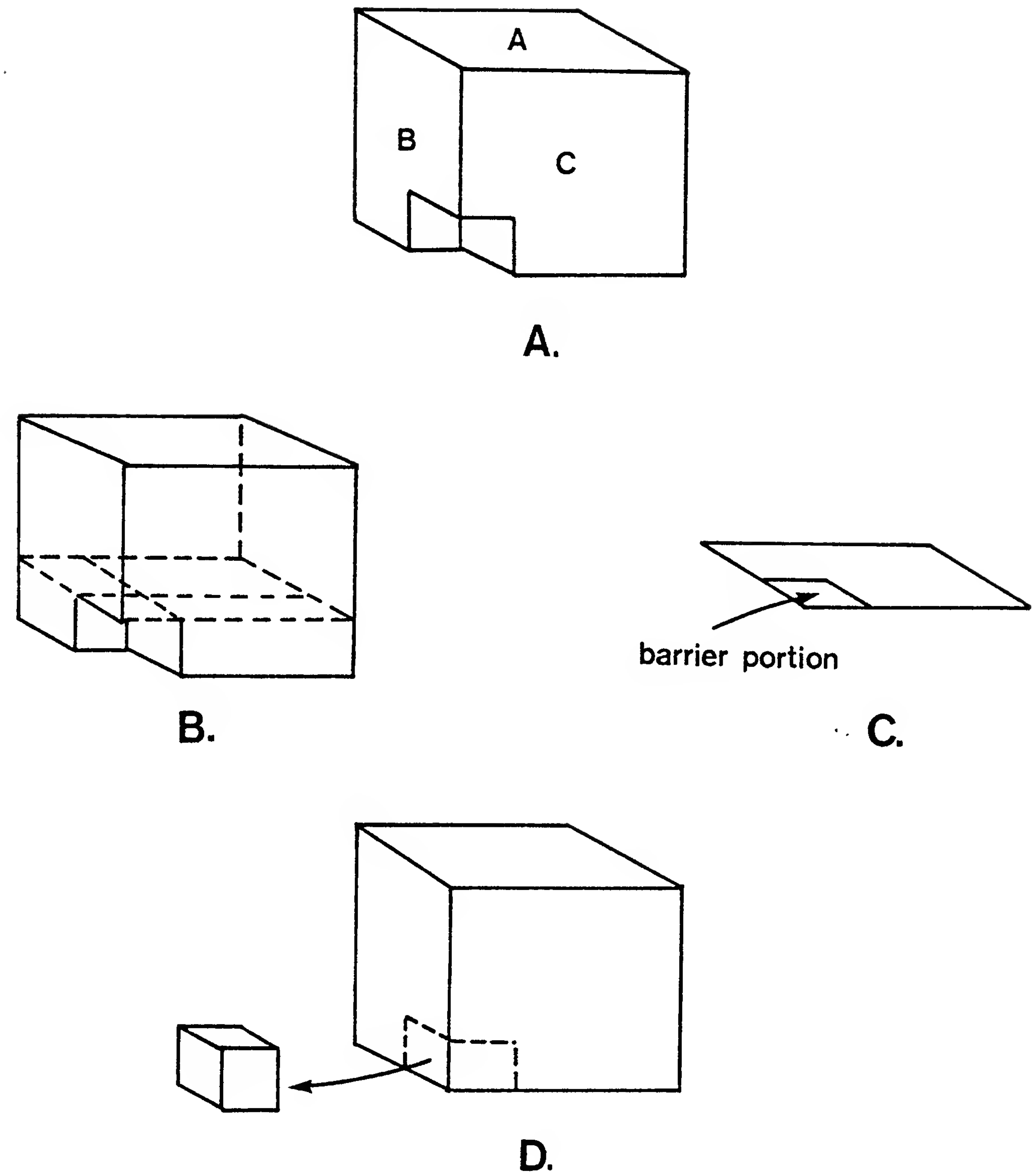


FIGURE 4.36.

figure 4.37B seems to lie outside the region, as if it were added on; hence it should be interpreted as a protrusion. It would appear contradictory to continue projecting A to yield an indentation if the barrier part of A looks like a protrusion. Instead, one should start the projection all over again with a modified region A' (figure 4.37C). Note that by retaining the added barrier portion lines in figure 4.37B, the segmentation routine could work on figure 4.37C to yield the two objects in figure 4.37D.

To distinguish these two cases, a simple definition for inside versus outside is offered. A barrier portion lies inside if, when it is removed from the cross section, the cross section becomes more complex; otherwise it lies outside. Thus removing the barrier portion in figure 4.36C makes the cross section more complex (an 8-sided region), while removing it in figure 4.37B makes the cross section simpler (a rectangle).

Exterior Barriers. Whether a barrier is interior or exterior depends on the direction in which it is approached. The objects in figures 4.35-7 are shown inverted in figures 4.38A-C respectively. The cross sections A at these cylinder ends see the barriers as exterior. Ideally the same descriptions should be obtained, and indeed the situations are treated correspondingly.

In figure 4.38A the barrier proportion is greater than 50% of the *combined* areas A+B, and so the cylinder formed thus far with A is segmented as a protrusion. Projection continues with B to yield the same description as in figure 4.35E.

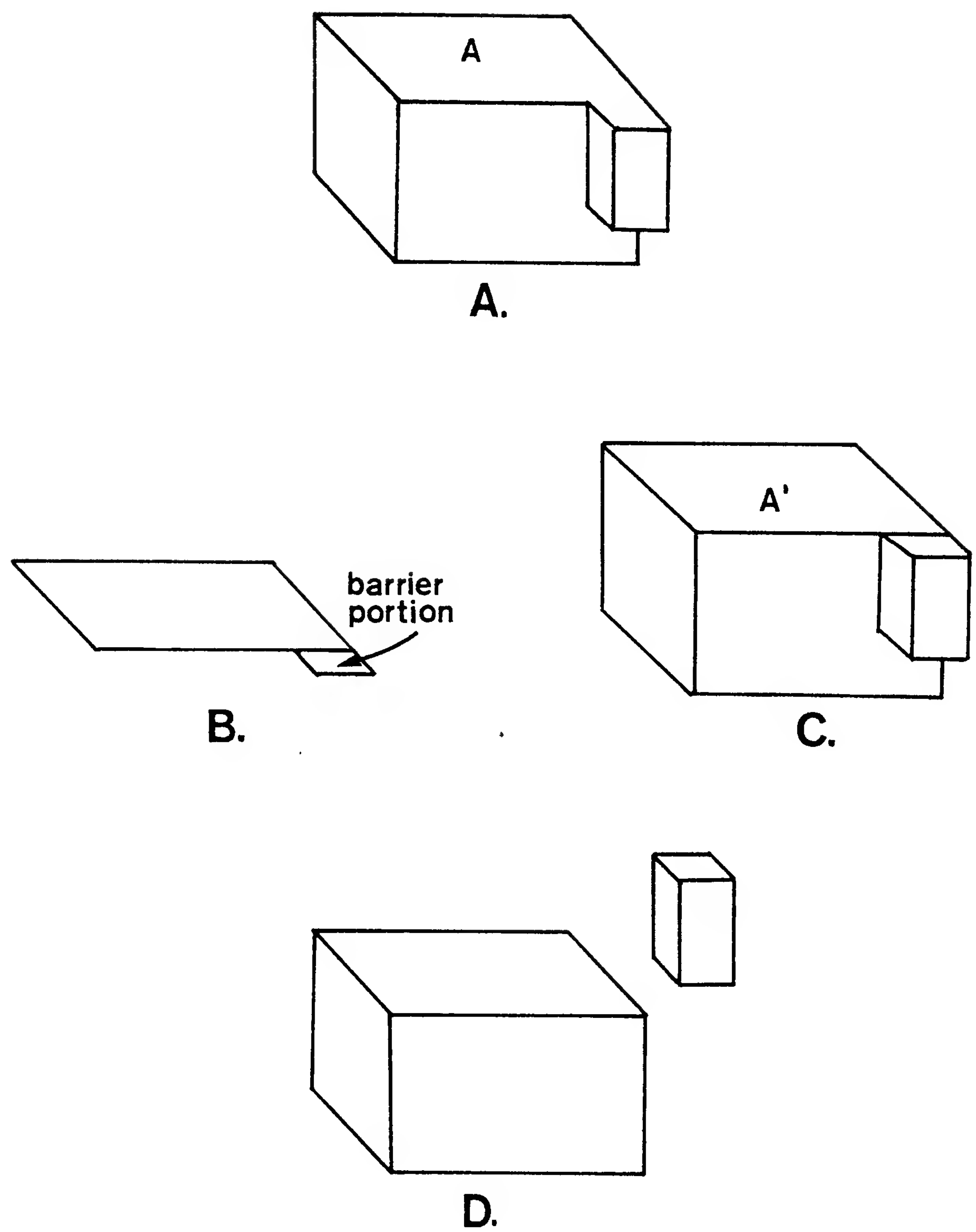


FIGURE 4.37.

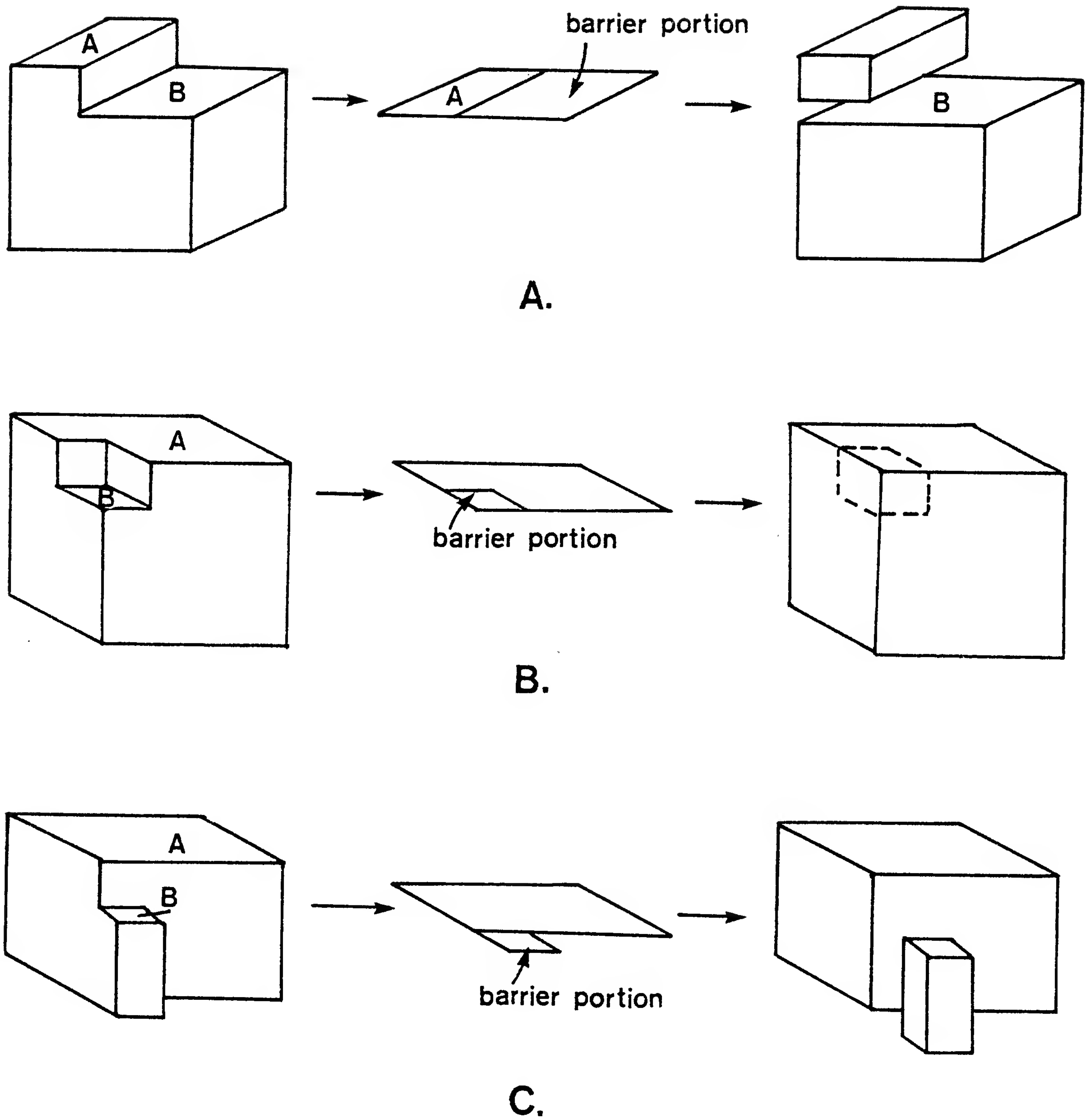

A
B
A
barrier portion
B
A.
A
B
barrier portion
B.
A
B
barrier portion
C.

FIGURE 4.38.

The barrier proportion is less than 50% of A+B in both figures 4.38B and C. In figure 4.38B a more complex cross section arises when B is substracted from A+B, and hence the barrier is interpreted as an indentation. The cross section is augmented in the continued projection by region B. In figure 4.38C a less complex cross section arises if B is subtracted from A+B, and so the barrier is segmented as a protrusion.

A complicating factor ignored here is the distance projected. If small, it tends to make an external barrier look like an indentation.

AN ILLUSTRATIVE DESCRIPTION OF AN OBJECT WITH SEVERAL TYPES OF BARRIERS

Returning to the object in figure 4.32, step 3 of the procedure calls for obtaining a description with region A and comparing it against B's. Continuing from figure 4.32B where a protrusion has been removed, region A when projected reaches an exterior barrier. This barrier is less than 50% of the combined cross section and barrier area, and its removal makes the augmented cross section less complex (figure 4.39A). Hence projection continues past it to yield a protrusion.

In obtaining the alternate description with B in figure 4.32A, an exterior barrier is first reached in the upper portion. Barrier removal results in a more complex region (figure 4.39B), and so B is augmented with the barrier region to become an L-shaped region. When continuing the projection of this augmented region, an interior barrier is encountered next (figure 4.39C). Since this barrier portion is less than 50% of the area, and since its removal makes the cross section less complex, the

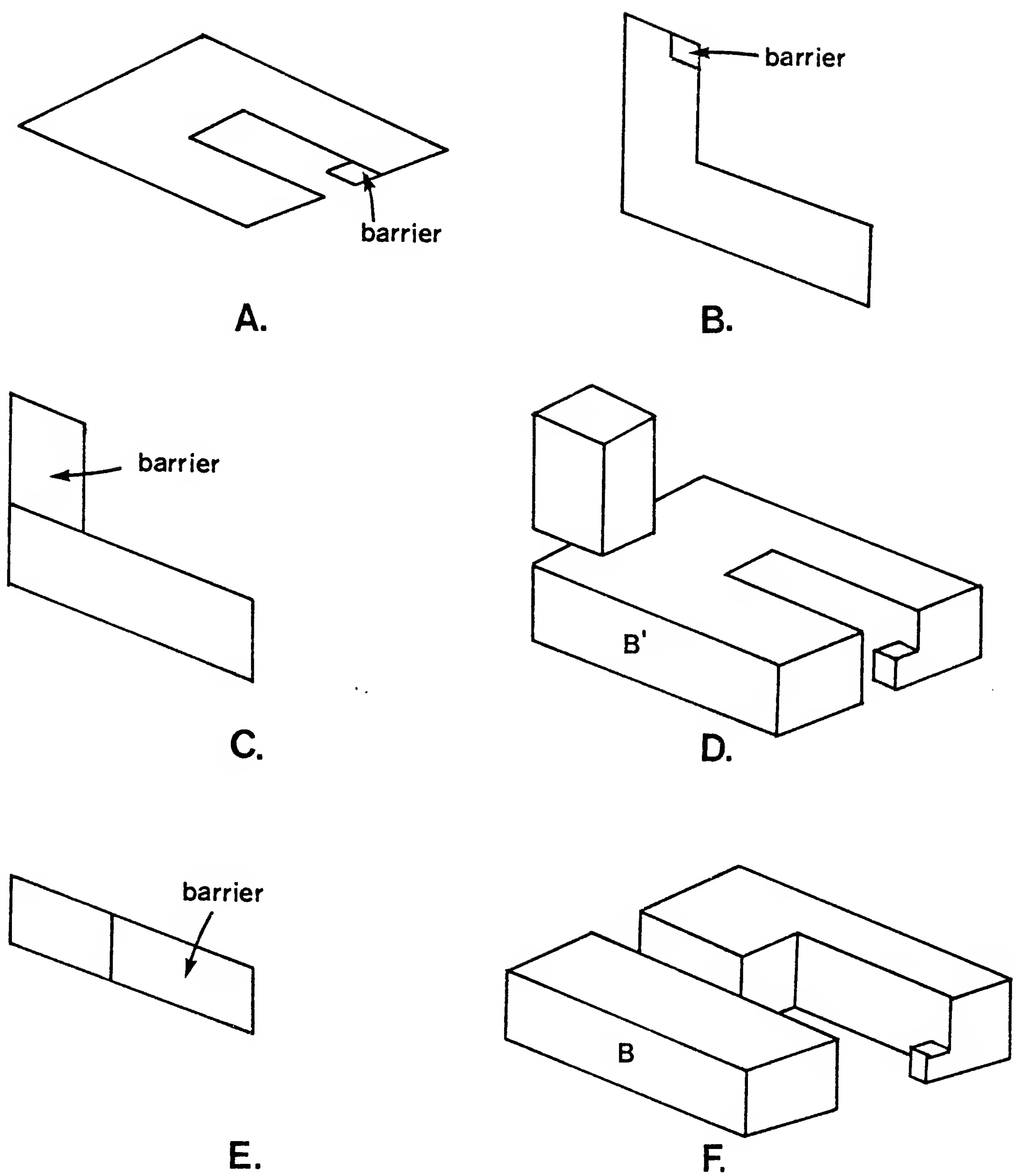


FIGURE 4.39.

original cross section is dissected and the barrier portion is segmented as a protrusion (figure 4.39D). The decremented area B' when projected now reaches an interior barrier, which is seen to comprise more than 50% of the area (figure 4.39E). Hence segmentation stops and produces a rather complicated protrusion (figure 4.39F).

Comparing the descriptions using A then B, it is clear A gives the simpler description.

4.7 Some Problems and ββßgestions

Smoothing. Small modifications greatly complicate the processing, as is evident from the preceding section. Every line that exists on an object must be taken into account, so as to either dismiss it as a minor modification or to recognize it as a major point of the description. Analogous to the curved object domain, some smoothing preprocess would be extremely useful to detect and simplify minor features. This would simplify enormously the tasks of choosing a cross section and carrying out the projection.

Some smoothing is done presently during the process of projection, resulting either in filling in indentations, removing protrusions, or altering the cross section. Perhaps this smoothing could be done more systematically during a pre-projection phase. Minor features might be indicated by short lines, irregular parts of regions, and protrusions on the contour. Various regions could be simplified beforehand by subjecting them to a prototype-modifier analysis, since this is also done frequently in the present procedure. In other words, a greater facility for jumping around to various parts of an object, sampling features along the way and hypothesizing their nature, is needed before the main descriptive process can be carried out.

Assembly Shapes. One problem not dealt with here is the shape of multiple object assemblies. For stacked alignment, the lines of alignment have almost nothing to do with shape, and could just as well be absent

without modifying shape. Thus the arch in figure 4.40A with lines removed still looks like an arch in figure 4.40B, although it is composed of one object instead of three. Protrusions are also very close to stacked alignment; by adding a line of alignment here or there the protrusion can be made to look like a separate aligned object. In fact the present procedure made use of this similarity in the segmentation of protrusions by adding appropriate lines of alignment and then applying the parsing procedure to them. The question is, given this similarity with respect to shape, why are they processed so differently? Since stacked assemblies and single objects can look exactly the same, should not they be processed in the same way? Maybe segmentation of aligned objects is a red herring; first, we should be concerned with overall shape, and maybe then we want to look for lines of alignment to determine composition. We should not allow these lines to result in a parsing strategy first.

Spurious Lines. Not considered in this approach are spurious lines, which seem to cause considerably more difficulty than missing lines. Shadow lines are one type of spurious line, which Waltz is able to account for by semantic limitations. Stray lines as might be introduced by a linefinder, however, cannot be treated in the same manner. Situations with spurious lines could be handled similar to those with missing lines, namely by hypothesizing a best object from erroneous ray data.

Some ways spurious lines may disturb scene analysis are by creating pseudo-rays at vertices, by blocking the path of a projection, or by

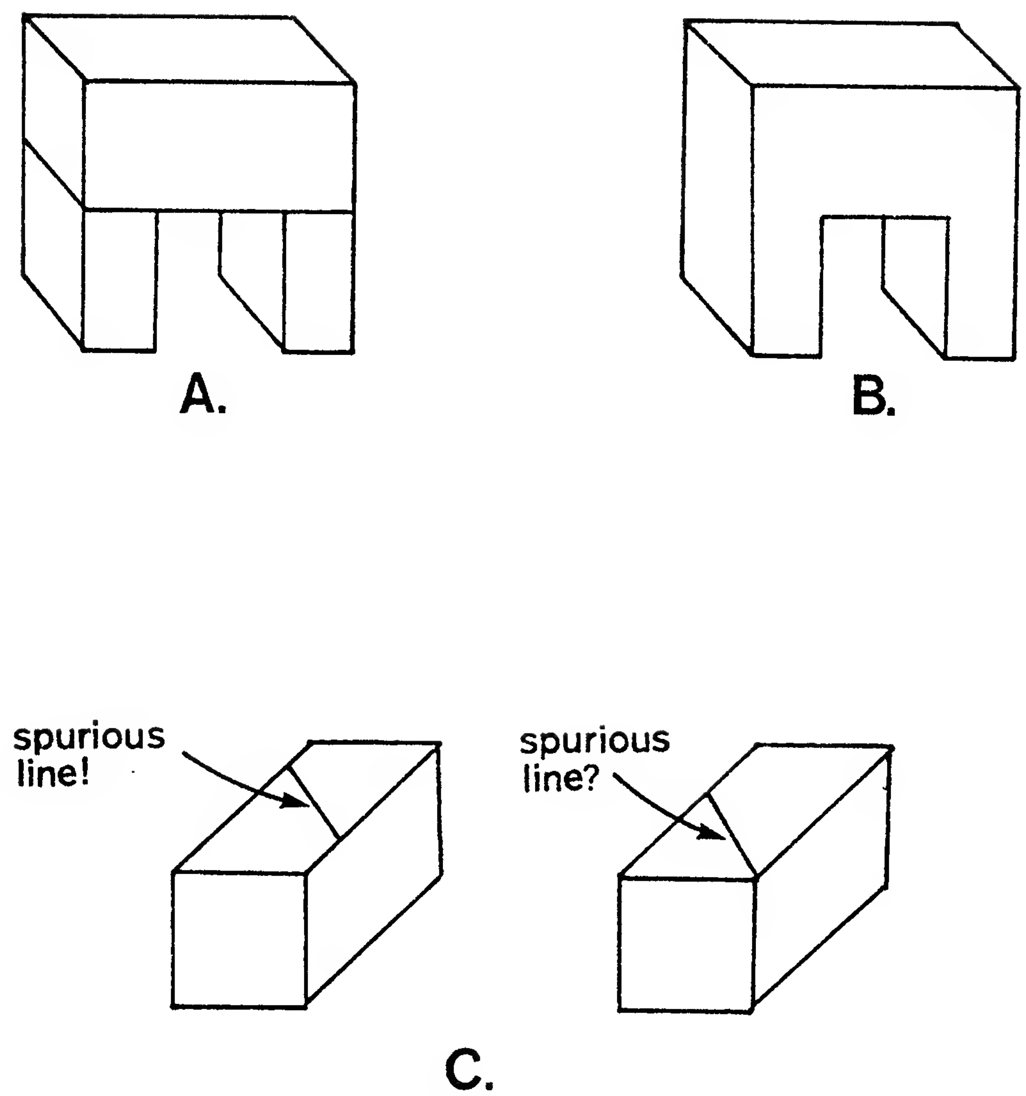


FIGURE 4.40.

dissecting a cross section. Pseudo-rays could be handled by selecting a subset that yields the nicest projection. A blocked projection could be continued past a spurious line in order to obtain a nicer object. Some spurious lines may appear as lines of alignment, and hence are hard to detect as such. Thus the line of alignment in figure 4.40C could represent the join of a wedge and trapezoidal block, or it could be a stray line in a rectangular block. The shape of the assembly, in any case, remains the same.

A final problem of the present approach is the built-in bias towards segmentation. Thus the first parsing of the object in figure 4.41 would be found but not the other two. Waltz finds all three parsings as well as an interpretation as an inseparable object. Whether this bias is a bug or feature is moot.

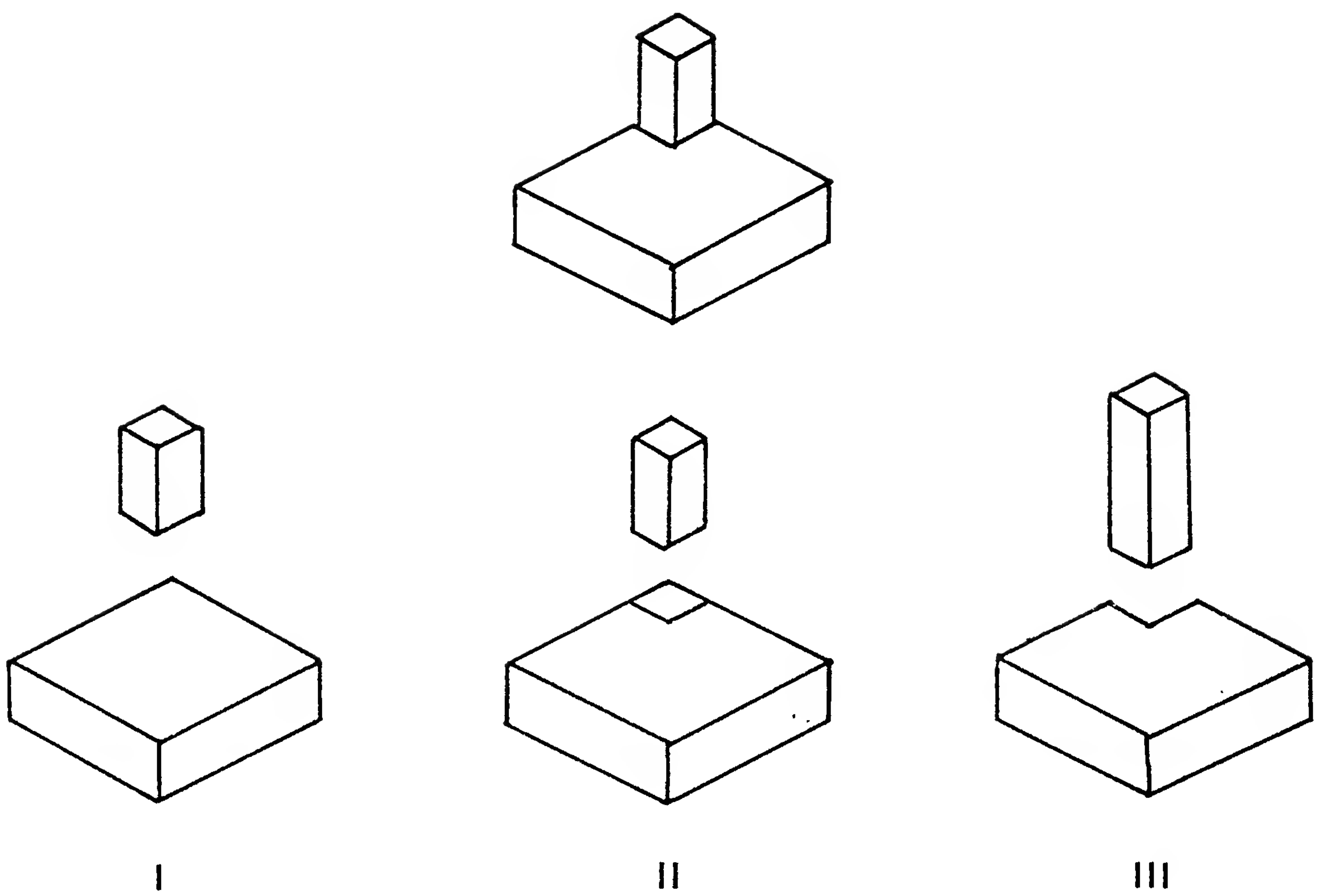


FIGURE 4.41.

CHAPTER 5 -- Concluding Remarks

This section concludes the thesis by relating my work to current research in AI, by discussing its generality and applicability to other domains, and by presenting suggestions for further work.

FRAMES AND DEBUGGING

The relation of prototypes and modifiers to frames and their slots has been discussed earlier. The frame for describing pottery borrows heavily from the frame for the human body, inasmuch as terms like foot, body, neck, lip, and mouth describe parts of a vase and their spatial relationship.

Frames provide a paradigm for approaching intelligent tasks, but they do not solve the tasks. Considerable thought must be given to what the descriptive elements of a frame are and to their relationships. This may in the final analysis be the central problem. If the present thesis is any indication, coming up with good descriptions is formidable.

Current research also focuses around the concept of debugging [Goldstein 1974 and Sussman 1973]. Modification assignment can be construed as debugging a simple prototype hypothesis to conform more exactly to an object shape. As Gombrich pointed out, a simple hypothesis is not more probably right, but rather more easily refuted and modified.

ARCHEOLOGISTS'S DESCRIPTIONS AS PROTOCOLS

Present work in medical diagnosis and past work by Newell and Simon in problem solving have relied on the use of protocols. The protocols are analyzed in order to deduce the structure of knowledge of the doctor in the first case and of the problem-solving subject in the second. Vision has been thought immune to language analysis [Rubin 1974], subject to study perhaps only by neurophysiology, by psychological experiments, or by the constructive approach of machine vision workers. What I have done in this thesis, however, is to analyze a sort of protocol: archeologists's vase descriptions.

The basis for using these protocols is a hypothesis that analyzing the structure of utterances about visual properties of objects such as shape says something about the structures inside a person's head. Archeologists are undoubtedly more expert at describing shape than ordinary persons, so that studying their descriptions corresponds to the current practice of interrogating experts about knowledge in their particular domain.

THE GENERALITY OF THE APPROACH

An indication of the generality of the approach is the similar treatment of the seemingly unrelated domains of pottery and polyhedra. Many natural objects are obvious cylinders. Prototype modification does not depend on generalized cylinders, but is a concept that applies to many object domains. Each object domain will possess unique prototypes and specific forms of modification to them.

How general are the specific descriptive terms developed for pottery? The shape descriptors are common everyday terms, and are applied to many sorts of objects besides vases. I offer as further proof of generality some observations on the use of the naming program by non-archeologists. This naming program can be caused to interrogate a person for vase descriptions rather than a stored data base. It directs to a person specific questions, such as "is the neck narrow?", to which the person replies only yes or no. People untrained in pottery description were asked to describe common objects like coke bottles, coffee cups, and jars to the program. In all cases they found the terms natural to their own usage, and had no difficulty in assigning meanings to a term; they readily decided, for example, whether a neck was narrow or not. One person chose to experiment with the program by describing a light bulb, and was himself surprised when the light bulb was aptly named a flask.

A simple cylinder or tube seems to be a starting point for many biological shapes. To understand the deviations from a simple cylinder is to understand the forces that form it.

> "Nature, like a glassblower, often starts with a simple tube. The stomach is an ill-blown tube, a bubble that has been rendered lopsided by a trammel or restraint along the side, such as to prevent a symmetrical expansion--such a trammel as is produced if the glassblower lets one side of his bubble get cold, and such as is actually present in the stomach itself in the form of a muscular gland. Nature does just what the glassblower does, and, we might even say, no more than he. For she can expand the tube here and narrow it there; thicken its walls or thin them; blow off a lateral offshoot or *caecal diverticulum*; bend the tube or twist and coil it; and infold or crimp its walls as, so to speak, she pleases." [Thompson p.1050]

Some biological organisms also look very much like pottery, as Thompson has remarked and from whom examples are taken in figure 5.1. Presumably these organisms could be described with pottery terms.

The pottery terms can be combined with polyhedral terms to describe objects such as telephones (figure 5.2). The body is a wedge-shaped simple cylinder, with a large indentation in the upper left corner and a slight truncation at the lower right. In the large indentation are two small U-shaped protrusions, which act as handle supports. The handle is a bow-shaped cylinder with rectangular cross section, joined at either end to two small bowls.

SUGGESTIONS FOR FURTHER WORK

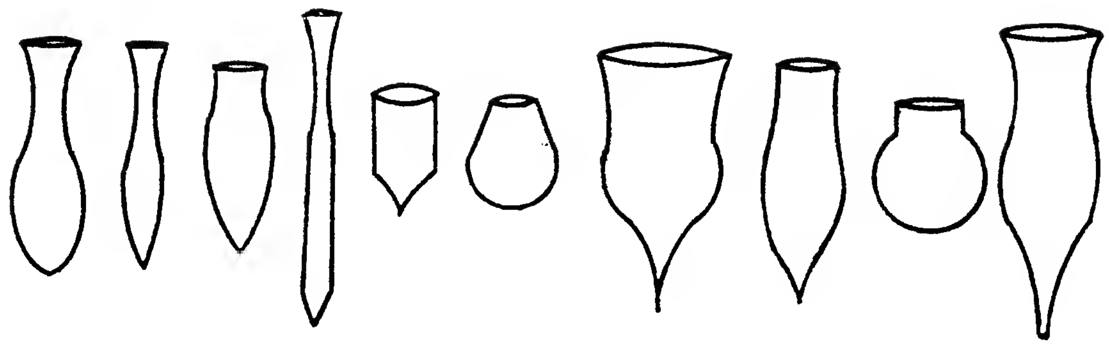
The following suggestions have presented themselves to me during the course of this thesis. They are either topics not treated or treated incompletely, or are suggestions for applying the descriptive method.

1. Multiple cylinder objects

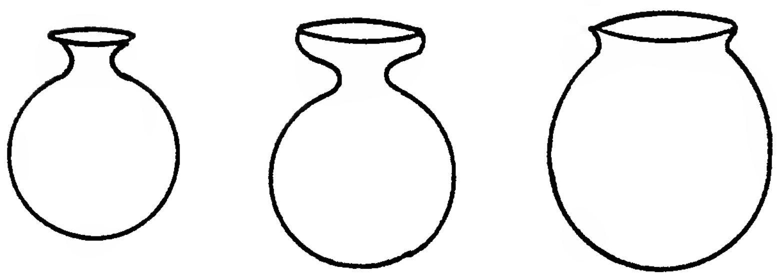
The major thrust of this thesis, both for polyhedra and for pottery, has been the description of single cylinders. It would be desirable to expand this work to multiple cylinder objects. The main difficulty is segmentation.

2. Handles

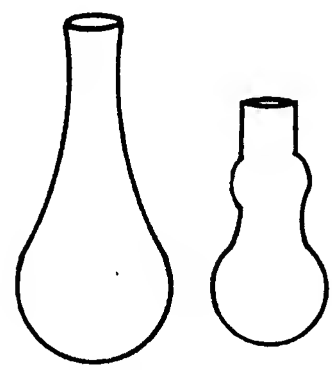
Descriptive terms for handles were presented, but these were not derived from some form of visual input. Actually a special case of suggestion 1, a good project would be to detect and segment handles in all



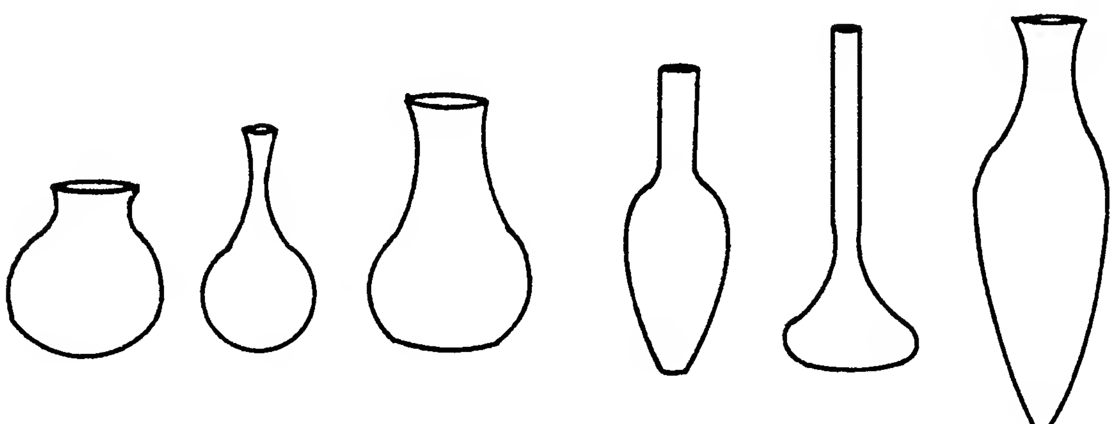
Various species of *Tintinnus*, *Dinobryon*, and *Codonella*.



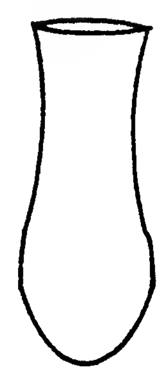
Flask-shaped shells or cysts.



Folliculina.



Various species of *Lagena*.



Vaginicola.

FIGURE 5.1. Some biological organisms show remarkable similarity to pottery, from *Thompson* (1952).

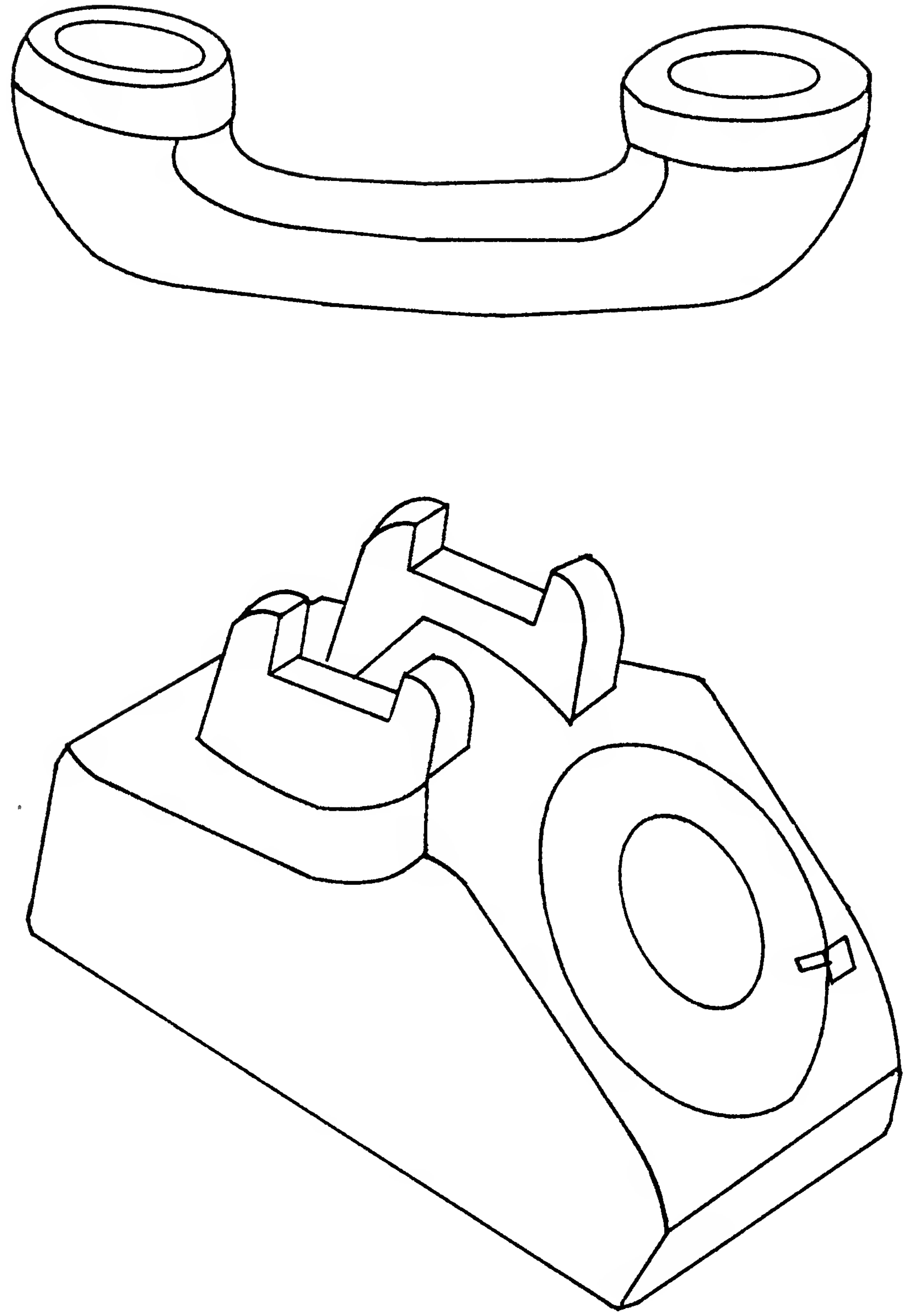


FIGURE 5.2.

sorts of positions relative to the viewer.

3. Noncircular cross sections

For the pottery domain it was assumed all cross sections were circular. A difficult problem is the detection of noncircular cross sections. What visual properties must be used? Texture? Shading? Highlight?

4. Smoothing

For both pottery and polyhedra, some effort was spent on smoothing minor irregularities to facilitate building a coarse description. Find better ways of smoothing. For polyhedra in particular, a preprocess smoothing would simplify object description, because small nicks greatly complicate the projections.

5. Shape primitives

Develop and implement a better set of shape descriptors. These are not necessarily the polished descriptors, but rather enumerate local features, such as corner, angle, ragged outline, and wavy. People in fact are better at describing local features than they are at building descriptions. They more readily compare objects, noting local differences, than they describe objects in isolation.

6. Form versus function

Description often depends on function of an object, on its material of construction, and on its method of construction. World knowledge about the uses and properties of objects must come into play in the descriptive process. This is amply demonstrated in the pottery domain, where the

malleability of clay reflects itself in plastic modifiers and where handle description depends on function. Investigate this dependence further.

REFERENCES

Agin, Gerald Jay [1972]: Representation and Description of Curved Objects. Memo AIM-173, Stanford AI Project.

Arnheim, Rudolph A. [1954]: *Art and Visual Perception.* Berkeley: University of California Press.

Attneave, Fred [1954]: Some Informational Aspects of Visual Perception. *Psych. Rev.* 61:183-193.

Attneave, Fred, and Arnoult, Malcolm D. [1956]: The Quantitative Study of Shape and Pattern Perception. *Psych. Bull.* 53:452-471.

Attneave, Fred [1957]: Transfer of Experience with a Class-Schemata to Identification-Learning of Patterns and Shapes. *J. Exp. Psych.* 54:81-88.

Attneave, Fred [1967]: Criteria for a Tenable Theory of Form Perception. *In: Models for the Perception of Speech and Visual Form.* Weiant Wathen-Dunn, ed. MIT Press.

Bartlett, Sir Frederick C. [1932]: *Remembering.* Cambridge University Press.

Binford, Thomas O. [1971]: Visual Perception by Computer. *IEEE Conference on Systems and Control.* Miami.

Birkhoff, George D. [1933]: *Aesthetic Measure.* Harvard University Press.

Blum, Harry [1964]: A Transform for Extracting New Descriptors of Shape. *In: Symposium on Models for Perception of Speech and Visual Form.* Boston.

Clowes, M. B. [1971]: On Seeing Things. *AI Journal.* 2:no. 1.

Coldstream, J. N. [1968]: *Greek Geometric Pottery: A Survey of Ten Local Styles and Their Chronology.* London: Methuen.

Cook, R. M. [1960]: *Greek Painted Pottery.* London: Methuen.

Falk, Gilbert [1970]: Computer Interpretation of Imperfect Line Data as a Three-Dimensional Scene. Memo AIM-132, Stanford AI Project.

Finin, Tim [1971]: Finding the Skeleton of a Brick. Vision Flash 19, MIT AI Laboratory.

Finin, Tim [1972]: A Vision Potpourri. Vision Flash 26, MIT AI Laboratory. Also in [Winston 1973].

Freiling, Michael J. [1973]: Functions and Frames in the Learning of Structures. Working Paper 58, MIT AI Laboratory.

Freuder, Eugene C. [1973]: Active Knowledge. Working Paper 53, MIT AI Laboratory.

Gabriel, Richard P. [1973]: Structured Descriptions. Vision Flash 48, MIT AI Laboratory.

Gardin, Jean-Claude [1967]: Methods for the Descriptive Analysis of Archeological Material. *American Antiquity*. 32:13-30.

Gardin, Jean-Claude [1972]: Etude d'un Systeme Integre d'Acquisition et d'Exploitation Automatique de Donnees Scientifiques. Marseilles: C.N.R.S., Centre d'Analyse Documentaire pour l'Archeologie.

Gluchoff, Alan [1973]: The Child's Conception of Three-Dimensional Objects. Term Paper for 6.544J (Minsky, Papert) at MIT, Spring 1973.

Goldstein, Ira [1974]: Understanding Simple Picture Programs. TR-294, MIT AI Laboratory.

Gombrich, E. H. [1965]: *Art and Illusion*. Princeton University Press.

Guzman, Adolpho [1967]: Some Aspects of Pattern Recognition by Computer. AI TR-224, MIT AI Laboratory.

Guzman, Adolpho [1968]: Computer Recognition of Three-Dimensional Objects in a Visual Scene. AI TR-228, MIT AI Laboratory.

Guzman, Adolpho [1970]: Analysis of Curved Line Drawings Using Context and Global Information. *In: Machine Intelligence 6*. Edinburgh University Press.

Hebb, D. O. [1949]: *Organization of Behavior*. New York: Wiley.

Hollerbach, John M. [1972a]: An Approach to Three-Dimensional Decomposition and Description of Polyhedra. Vision Flash 31, MIT AI Laboratory.

Hollerbach, John M. [1972b]: The Projective Approach to Object Description. Vision Flash 37, MIT AI Laboratory.

Horn, Berthold K.P. [1970]: Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View. AI TR-232, MIT AI

Laboratory.

Huffman, David [1971]: Impossible Objects as Nonsense Sentences. *In: Machine Intelligence 6.* Edinburgh University Press.

Krakauer, Lawrence J. [1971]: Computer Analysis of Visual Properties of Curved Objects. AI TR-82, MIT AI Laboratory.

Kuhlmann, F. [1906]: On the Analysis of Memory Consciousness: A Study in the Mental Imagery and Memory of Meaningless Visual Forms. *Psych. Rev.* 13:316-348.

Lacy, A. D. [1967]: *Greek Pottery in the Bronze Age.* London: Methuen.

Marr, D., and Hewitt, C. [1973]: VIDEO ERGO SCIO: An Essay on Some Things We Would Like a Vision System to Know. Working Paper 60, MIT AI Laboratory.

Maruyama, Kiyoshi [1972]: A Study of Visual Shape Perception. UIUCDCS-R-72-533, Department of Computer Science, University of Illinois at Urbana-Champaign.

Minsky, Marvin L. [1974]: A Framework for Representing Knowledge. AI Memo 306, MIT AI Laboratory.

Nevatia, Ramakant [1974]: Structured Descriptions of Complex Curved Objects for Recognition and Visual Memory. Memo AIM-250, Stanford AI Project.

Noble, Joseph V. [1965]: *The Techniques of Painted Attic Pottery.* New York: Watson-Guptill Publications.

Posner, Michael I. and Keele, Steven W. [1968]: On the Genesis of Abstract Ideas. *J. Exp. Psych.* 77:353-363.

Rattner, Martin H. [1970]: Extending Guzman's See Program. AI Memo 204, MIT AI Laboratory.

Richter, Gisela M.A., and Milne, Marjorie J. [1973]: *Shapes and Names of Athenian Vases.* Washington: Mcgrath Publishing Co.

Roberts, L. [1965]: Machine Perception of Three-dimensional Solids. *In: Optical and Electrooptical Information Processing.* Tippett, J. T., et al (ed.). MIT Press.

Rosch, Eleanor [1973]: On the Internal Structure of Perceptual and Semantic Categories. *In: Cognitive Development and the Acquisition of Language.* Moore, T.E. (ed.). New York: Academic Press.

Rubin, Ann D. [1974]: Artificial Intelligence Approaches to Medical Diagnosis. Working Paper 65, MIT AI Laboratory.

Shirai, Yoshiaki [1972]: A Heterarchical Program for Recognition of Polyhedra. AI Memo 263, MIT AI Laboratory. Also in [Winston 1973].

Sussman, Gerry J. [1973]: A Computational Model of Skill Acquisition. TR-297, MIT AI Laboratory.

Thompson, D'Arcy [1952]: *On Growth and Form.* Cambridge University Press.

Waltz, David L. [1971]: Understanding Scenes with Shadows. Vision Flash 21, MIT AI Laboratory.

Waltz, David L. [1972]: Generating Semantic Descriptions from Drawings of Scenes with Shadows. AI TR-271, MIT AI Laboratory. Also in [Winston 1973].

Warnock, John E. [1969]: A Hidden Surface Algorithm for Computer Generated Halftone Pictures. Computer Science Technical Report No. 4-15, University of Utah.

Warren, Peter [1969]: *Minoan Stone Vases.* London: Cambridge University Press.

Whallon, Robert, Jr. [1972]: The Computer in Archeology: A Critical Survey. *Computers and the Humanities.* 7:29-45.

Winston, Patrick H. [1970]: Learning Structural Descriptions from Examples. AI TR-231, MIT AI Laboratory.

Winston, Patrick H., ed. [1973]: Progress in Vision and Robotics. AI TR-281, MIT AI Laboratory.

Woodworth, R.S. [1938]: *Experimental Psychology.* New York: Holt.

Zusne, Leonard [1970]: *Visual Perception of Form.* New York: Academic Press.